NAS12-2014

FINAL REPORT

STUDIES AND ANALYSES OF AUTOMATIC COMMAND CONTROL COMMUNICATIONS AND APPLICATIONS OF TIME REFERENCE SYSTEMS FOR AUTOMATIC FLIGHT MANAGEMENT

By H. C. Salwen, S. M. Sussman, C. H. Boardman

26 January 1969

Prepared under Contract No. NAS12-2014

National Aeronautics and Space Administration
Electronics Research Center
Cambridge, Massachusetts 02139

ADVANCED COMMUNICATIONS INFORMATION MANAGEMENT

RESEARCH • DEVELOPMENT • ENGINEERING

ADCOM

A TELEDYNE COMPANY

Mr. John Loria
Technical Monitor
NAS12-2014
575 Technology Square
Cambridge, Massachusetts 02139

Requests for copies of this report should be referred to:

National Aeronautics and Space Administration
Electronics Research Center
Cambridge, Massachusetts 02139

NAS12-2014

FINAL REPORT FOR STUDIES AND ANALYSES OF AUTOMATIC COMMAND CONTROL COMMUNICATIONS AND APPLICATIONS OF TIME REFERENCE SYSTEMS FOR AUTOMATIC FLIGHT MANAGEMENT

By H.C. Salwen
S.M. Sussman
C.H. Boardman

26 January 1969

Prepared under Contract No. NAS12-2014 by

ADCOM
A Teledyne Company
808 Memorial Drive
Cambridge, Massachusetts 02139

For

National Aeronautics and Space Administration
Electronics Research Center
Cambridge, Massachusetts 02139

G-140F

CONTENTS

	Page
INTRODUCTION	1
Object of Study Program	1
Summary of the Report	1
SYSTEMS CONSIDERATIONS	5
Ground/Air/Ground Direct	5
Radio Navigation Aids	5
Radar Surveillance	6
Direct Air/Ground Communication	6
Air-to-Air Direct	8
Ground-to-Air via Satellite	8
Air-to-Ground via Satellite	10
Concluding Remarks	12
COMBINED RANGING AND COMMUNICATION SIGNALS	13
Pseudorandom Range Code Systems	13
Range Code Polarity Reversal	21
Ranging from Bit Synchronization	22
Combined Harmonic Ranging and Data Transmission	28
Summary	32
NAVIGATION USING TIME REFERENCE SYSTEMS	35
Background	35
Utilization of Surfaces in Trajectory Situations	37
Factors Affecting GDOP	37
Equivalence of Solutions in the Just-Determined Case	42
The Least-Squares Data Reduction Technique	45
Least-Squares Estimates of Parameters Based On a Series of Similar Measurements	45
Least-Squares Combining of Data From a Variety of Measurements	50
Investigation of GDOP Phenomena Using the Least-Squares Data Reduction Technique	54
GDOP in Hyperbolic and Augmented Hyperbolic Navigation Systems	54
Results of the Hyperbolic Navigation Investigation	56
GDOP in a Multilateration Altimeter	61
Results of Multilateration Altimeter Analysis	65

CONTENTS (Cont'd)

Page

AVAILABILITY OF STABLE SOURCES FOR TIME REFERENCE SYSTEMS 69

Characteristics of Frequency Standards 69
Applicability to Passive Navigation Systems 72

Appendix A
MULTIPATH 75

Appendix B
PSEUDORANDOM RANGING ERROR 107

Appendix C
BIT SYNCHRONIZATION TIMING ERROR 111

Appendix D
NEW TECHNOLOGY 115

ILLUSTRATIONS

Figure		Page
1	Autocorrelation of Pseudorandom Ranging Code	13
2	Typical (Baseband) Receiver Structure for Pseudorandom Code Ranging	14
3	Acquisition Performance of Pseudorandom Ranging System	16
4	Bound Involving Acquisition Time, Code Length, and SNR (Equation 3)	17
5	Control Characteristic in the Tracking Mode	18
6	Ranging Accuracy as a Function of Bit Duration and SNR (Assuming $T_{int} = 10^{-2}$ sec in Eq. 4)	19
7	Relationship Between Maximum Unambiguous Range, Code Length, and Bit Period	20
8	Trade-off Between Data Rate and Maximum Unambiguous Range	21
9	Complement Correlation Function of 00001010111011000111110011010 01	23
10	Error Performance of Various Binary Modulation Methods	24
11	Receiver Structure Including Bit Sync Extraction	26
12	Ranging Accuracy of Bit Sync System	27
13	Spectra of Modulating Waveforms	29
14	Spectrum of Data Modulation	31
15	Split-Phase Waveforms and Spectrum	33
16	Intersection of Surfaces Defines Position	36
17	Geometry of Hyperbola and Ellipse	39
18	Uncertainty in Location of Surface	41
19	Intersection of Circle, Hyperbola, Ellipse	43
20	Simple Position Determination Problem	45
21	Arbitrary Range-Angle Measurement (in 2 Dimensions)	51
22	Orientation of Computer Printout Coordinate System with Station Configuration III Superimposed	57

ILLUSTRATIONS (Cont'd)

Figure			Page
23	a	Run #1	57
	b	Run #3	57
24	a	Run #1	59
	b	Run #7	59
25	a	Run #3	59
	b	Run #9	59
26	a	Run #3	60
	b	Run #15	60
27	a	Run #1	60
	b	Run #19	60
28	a	Run #3	61
	b	Run #21	61
29	a	Run #21	62
	b	Run #27	62
	c	Run #33	62
	d	Run #39	62
30	a	Run #19	63
	b	Run #25	63
	c	Run #31	63
	d	Run #37	63
31		Format of the Computer Printout	64
32		Position Errors as a Function of Position for a Multilateration Altimeter	66, 67
33		Frequency Standard Stability	70
34		Normalized Phase Departure vs. Observation Time for Various Frequency Standards	71
A-1		Practical Site Reauirements for Satisfactory Performance Natural Terrain Criteria	77
A-2		Geometrical Configuration of Station, Reflector and Aircraft	82
A-3		Phasor Diagram Representation of the Sum of Direct and Reflected Path Contributions to the AM Tone in Conventional VOR	85
A-4		Variation of a'/a with $\omega_c t_d$ for Various Values of $\underline{a}$	87

SECTION 1

INTRODUCTION

This document constitutes the Final Report prepared by ADCOM, a Teledyne Company, for NASA Electronics Research Center under Contract No.NAS12-2014 entitled "Studies and Analyses of Automatic Command and Control Communication and Applications of Time Reference Systems for Automatic Flight Management."

Object of Study Program

The objective of this study is set forth in the contract Statement of Work as

(a) Evolve a unified A/G and G/A automatic data communication, command and control link,

(b) Establish the basic design guidelines and requirements for signal management, signal detection and processing, and

(c) Evaluate the impact, technologically and performancewise, of precision time references on VOR, radar and Loran C systems as applied to air navigation.

The effort is divided into three items of investigation which may be summarized as follows: Item 1 is devoted to an overall review of the system functions, avionics and operational requirements of a unified A/G and G/A communication/navigation link for the purpose of establishing guidelines for the design of such a link. Item 2 is concerned with ways in which data links and time reference systems can be applied to automatic navigation and flight management. Item 3 is to determine the requirements imposed on time reference systems by these applications and determine the availability of time reference systems which meet these requirements.

Summary of the Report

The material in this report bears the following relation to the items in the Work Statement: Sections 2 and 3 plus the appendices pertain to Item 1, Section 4 covers Item 2 and Section 5 responds to Item 3.

Section 2 reviews the flow of information and signals on the various links to and from the aircraft. It is noted that the great disparity in carrier frequencies of existing radio navigation aids and communication links precludes their integration into a unified system. A future unified system is discussed that could incorporate satellite communication/navigation over unpopulated

areas, satellite navigation and direct or satellite communication (depending on satellite capacity) over populated regions, direct radiating landing aids, and air-to-air direct communication for special purposes. The data flow on ground-to-air links is described including ranging signals for navigation and various categories of communication traffic. It is recommended that ranging signals not be relayed on the air-to-ground link but that air traffic surveillance requirements be met by automatic position reporting in response to interrogations. Multiple access through the satellite by many aircraft is examined and it is concluded that frequency division channels allocated from a ground center are the preferred mode of operation.

In Section 3 several combined communication/ranging modulation schemes are described and analyzed. Trade-offs among ranging accuracy, data rate, signal-to-noise ratio and other pertinent system parameters are presented for each configuration. The transmission functions outlined in Item 1 of the Work Statement encompass a wide range of data rate requirements from roughly 10^6 bits per frame (i.e., 3×10^7 bits per sec at 30 frames per sec) for TV transmission down to roughly 100 bits per sec for flight management data. RF power requirements as a function of data rate and communication performance are presented in Section 3 for the various unified Comm/Nav modulation designs discussed. Potential malfunction due to interference of the data transmission function with the ranging navigation function within the system is analyzed and discussed. The power necessary to provide adequate ranging performance, as dictated by navigation requirements, is defined. It is found that pseudorandom ranging codes are severely limited in data capacity. On the other hand, harmonic ranging signals appear to be the most suitable for adaptation to carry data. When data rates are sufficiently high the bit stream itself can be used for ranging, thereby eliminating the need for separate ranging signals.

In response to Item 2 of the Work Statement the usefulness and applicability of precision time reference systems to various aircraft functions were analyzed. Preliminary evaluation showed that time reference systems with accuracy on the order of 1 μsec or better would find application in the area of aircraft navigation and perhaps communication.

Constraints imposed on time reference system requirement by the communication function are much less severe relative to the navigation constraints. In fact, the utility of time reference systems in the communication process, e.g., for synchronization in time division multiple access, depends on the establishment of accurate aircraft position data and/or accurate estimates of the propagation delay from the transmitting station to the aircraft and vice versa. Thus, it was concluded that application of such time reference systems to aircraft communication problems would be productive if and only if the position of the aircraft was known with an

accuracy commensurate with the time reference system accuracy. For these reasons, the utilization and effectiveness of on-board time reference systems for aircraft navigation are investigated in detail. Section 4 details the work performed and the results obtained during this phase of the study. Specifically, Section 4 is concerned with the characterization of errors in position determination associated with such techniques as LORAN-C, OMEGA, and DME/DME. Particular emphasis is placed on evaluating the utility of additional one-way passive ranging measurements obtained with the aid of an on-board stable time reference. The method of least-squares data reduction is derived and discussed. Examples are presented in which the least-squares method is used to determine errors in position location as a function of position, i.e., GDOP errors. The performance improvement obtainable if one-way ranging is accomplished with the aid of an on-board stable time reference system is investigated. In addition, the use of such techniques and systems for altimetry is investigated in detail.

It is found that time reference systems can be used to advantage in certain airborne navigation systems when such systems are under conditions of unfavorable geometry. Successful implementation of time reference systems in this application requires a calibration cycle which is repeated often enough so that the absolute timing error does not exceed 50-100 nsec.

The availability of frequency sources and time reference systems meeting the requirements of the navigation applications of Section 4 is considered in response to Item 3 of the Work Statement. The analyses performed and the results obtained through this effort are presented in Section 5. In particular, Section 5 discusses the state-of-the-art of stable sources for time reference systems. The relationship between typical stable oscillator specifications and time reference system specifications are established. Section 5 concludes with a discussion of the availability of stable sources for time reference systems which would fulfill the requirements outlined in Section 4 for aircraft navigation. It is found that a state-of-the-art quartz standard source would meet the timing accuracy requirements stated above if it was accurately recalibrated at 1 hour intervals. A state-of-the-art cesium standard would have to be recalibrated daily.

SECTION 2

SYSTEMS CONSIDERATIONS

Before going into some of the detailed guidelines for designing a unified RF system for aircraft communication and navigation, it is appropriate to review the flow of information and signals on the various links to and from the aircraft. For purposes of discussing a unified system these links can be classified into the following categories:

1) Ground/air/ground direct.
2) Air-to-air direct.
3) Ground-to-air via satellite.
4) Air-to-ground via satellite.

Ground/Air/Ground Direct

Radio Navigation Aids

Under this heading we can identify immediately a great variety of existing radio navigation aids ranging in frequency from VLF to microwaves. Among these systems are the long range navigation systems such as LORAN-C/D and OMEGA; the line-of-sight systems such as VOR and DME, and landing aids such as the ILS glide slope and localizer. Each of these systems was designed to serve a specific purpose and its carrier frequency was selected to correspond to certain desirable propagation phenomena or antenna patterns.

Most of the existing systems function with passive aircraft, i.e., the aircraft derives navigation information from received signals without radiating any signals of its own. On the other hand, a system such as DME requires active aircraft participation in that the aircraft interrogates a ground transponder.

Combining the existing systems into a unified system is obviously impractical in view of the disparity of carrier frequencies. Furthermore, a unified system should provide service over oceans as well as over populated land masses, preferably using the same aircraft components. A satellite system is clearly indicated for over-ocean flights and the same system could serve continental routes. For navigation purposes the only stipulation is adequate satellite coverage assuming passive aircraft and hence unlimited capacity for handling multiple users.

For communication the capacity question must be examined since the number of overland flights that must be served simultaneously may exceed a reasonably structured satellite communication system.

Whereas the possibility exists that satellite based navigation aids may eventually replace ground-based aids for enroute and area navigation, it is unlikely that landing aids would also be supplanted. Direct ground-to-air emission of landing aid signals remains the most attractive method for providing the necessary close-in accurate coverage. This, however, does not exclude the possibility of a high degree of commonality between the satellite navigation system and the landing aid. Both schemes can be based on the principles of time measurements and the establishment of circular or hyperbolic lines of position. A preliminary analysis of a ground-based time reference navigation/landing aid is discussed in Section 4.

A major problem with present direct navigation and landing aid systems is the perturbation of radiated field by reflections from nearby objects. Numerous failings of the existing systems have been attributed to the presence of multipath. An analysis of bearing errors in the VOR system due to multipath is presented in Appendix A. Appendix A discusses the problem of identifying sources of multipath disturbance. Then, detailed analyses of the effect of multipath on the various VOR implementations are presented. It is shown that an improvement in performance in the presence of multipath is obtainable via careful modulation design. Similar investigations must be undertaken for any proposed direct-radiating landing aid.

Radar Surveillance

An important direct ground-to-air link is radar surveillance with or without aircraft transponders. At present radar constitutes the primary independent surveillance capability of the Air Traffic Control system. Independence of aircraft computed position is a desirable feature of an ATC surveillance system. However, skin-tracking radar returns are the only truly independent observations and this will be compromised by the planned use of a coded transponder which identifies the aircraft and gives its altitude. It is no great step from the transponder to an automatic read-out of the on-board navigation system in response to an interrogation from ATC. It is a matter of the quantity and reliability of the equipment which interacts with the surveillance system. Further comments regarding automatic position reporting in response to interrogations are included in the section entitled "Air-to-Ground via Satellite" on page 10.

Direct Air/Ground Communication

Direct voice communication between ground and aircraft currently takes place on VHF radio links. With minor exceptions communication

concerning flight control, aircraft operations, meteorology, etc., is carried on these voice links. Some experimental use of digital data links on the same RF channels has been reported.* The aircraft industry sees great hope in data links and automated procedures to relieve the communication workload on personnel both in the air and on the ground. A recommended set of standards for a Universal Air-Ground Digital Communication System has been prepared by the Radio Technical Commission for Aeronautics.

If the satellite capacity and coverage is adequate there is fundamentally no need for direct communication between aircraft and ground. In fact, communication via satellite would obviate one of the major nuisances of the VHF system, viz., the recurring need for frequency changes. However, practical considerations of achievable satellite capacity (numbers of channels) and user equipment combine to ensure that direct communication will retain an important role in the foreseeable future. This raises the possibility of a dual mode system capable of direct communication or via satellite. Several important issues must be resolved before such a system could become a reality.

At the outset the matter of carrier frequency must be settled. The trend in satellite navigation and communication is toward frequencies higher than those presently employed by aircraft for these purposes. The preferred allocation is in a band near 1.6 GHz. With this choice the satellite RF link would be incompatible with the existing air-ground communication in the 118-136 MHz band. Retention of the present VHF ground stations would require, at the minimum, dual antennas and receiver front-ends for the aircraft. Moreover, the present AM voice communication format would be incompatible with the more advanced unified system signal design. The advent of data links with or without satellite communication will certainly require major modification of present operations and signaling practices. If a communication/navigation satellite system is under development in parallel with an air-ground direct data link, it is appropriate to consider making these two systems compatible in RF carrier frequency as well as modulation format. Eventual transition to the higher frequency for all communication could then take place for appropriately equipped aircraft, while others continued to employ the present VHF voice links.

If a direct ground-to-air system and a satellite system are to share the same aircraft receivers, the problem of mutual interference must be resolved. This will require thorough investigation of signal strengths and

---

* B.F.McLeod, "Digital Communication," RCTA Annual Assembly Meeting, September, 1968.

appropriate frequency channeling. The antenna patterns demanded by the two links are, of course, different. An earth-oriented as well as one or more space-oriented antennas will have to be installed onboard an aircraft with dual mode capability.

Air-to-Air Direct

Direct links between aircraft play an important role in collision avoidance and in search-and-rescue. At least two schemes are under consideration for collision avoidance: an optical/infrared Pilot Warning Indicator and a Time-Frequency Collision Avoidance System. The former is based on the detection of flashing Xenon lights and is clearly not unifiable with a satellite RF system.

Collision avoidance based on synchronized clocks aboard participating aircraft has received considerable attention and a rather detailed system configuration has been worked out. The suggested carrier frequency is in the 1.55 to 1.6 GHz band, nearly coincident with the preferred satellite link frequency. Thus antennas and possibly portions of the receiver front end could be shared among a CAS and a satellite system. However, mutual interference poses a serious problem which must be resolved before an even partially unified system of satellite, air-to-air, and direct air/ground links can be developed.

Search and rescue operations will also require direct communication as well as relative navigation between rescuer and distressed aircraft. Under emergency conditions the unified system can revert to a special mode for transmission of position location information via the satellite or by line-of-sight propagation to a rescue craft. Increased accuracy for relative navigation is possible by relaying raw measurements from distress craft rather than computed position. A properly equipped rescue aircraft can then fly to where its observations match those of the distress craft. Thereby inaccuracies introduced by calibration constants and similar error sources can be minimized.

Ground-to-Air via Satellite

In discussing the satellite links we separate the ground-to-air and the air-to-ground links and further subdivide each of these into to- and from-satellite portions. This breakdown is necessary because each portion has its distinct features and requirements.

As presently conceived, navigation satellite systems will consist of a cluster of four or more satellites at synchronous altitudes. Each satellite will emit a ranging signal suitable for precision measurement of time of

arrival at a user aircraft. The ranging signals are synchronized among the satellites so that the user can determine his position by measuring the relative time of arrival of signals from different satellites. The ranging signals can be generated at each satellite. However, for ease of inter-satellite synchronization and because of the need for communications and ranging, the signals are preferably generated at the ground with the satellite acting as a relay.

The ground-to-satellite link will utilize a large ground station and can accommodate a convenient carrier frequency (5 GHz or greater) without the logistics problems associated with the aircraft terminals. The capacity of this link is not critical in view of the large antennas and high power transmitters available at ground stations. Except for a change in carrier frequency, the downlink signals from the satellite to the aircraft will be the same as the uplink from ground to satellite.

The signals received at the aircraft will consist of the aforementioned ranging signals combined with the necessary communication traffic. Various schemes for providing a combined modulation format are discussed in Section 3. The parameters of the ranging signals are defined primarily by the desired range accuracy and ambiguity resolution. These parameters are not too difficult to establish, but the communications load is another matter.

The required capacity depends on such factors as the number of aircraft served by a particular satellite, the class of service, the duration and frequency of contacts and the tolerable queuing times for users awaiting service. Specific data should be ascertained from appropriate sources on the communication needs of a future system.

The communications traffic that would be carried on the ground-to-air link is comprised in part of broadcast-type of transmissions which are received by all aircraft in a given coverage area. Included here are:

Meteorological Data

Radiation Warnings

Search and Rescue Alerts and Other Emergencies

Interrogations and Channel Assignments

Entertainment

The first three items are self-explanatory. Weather data would be available to any aircraft on a routine schedule by selecting the channel which carries this information. Emergency messages would be broadcast so as to override other communications and to assure reception by all intended receivers.

Interrogations and Channel Assignments are low duty cycle communications whose purpose is to establish links between specific aircraft and ground terminals. Each aircraft listens on this channel for a call bearing its address code and a channel assignment for the return link. Automatic recognition circuitry would respond to the address code and either alert the flight crew or automatically switch on the indicated return or two-way link. In this manner various classes of information flow can be initiated by the ground terminal including position reports, airline flight management data, and passenger communications.

The final category of ground-to-air broadcast-type of communication is entertainment for passengers. This is a relatively low priority item and should not cause a major cost increment in a unified communication/navigation system. The main reason for this statement is the small size of the audience which makes entertainment broadcasts to aircraft unattractive for both commercial and airline interests, not to mention government subsidies.

In addition to the broadcast mode there will be specific ground-to-air communication to one aircraft only. This will consist of air traffic control instructions, company communication with the flight crew, and individual passenger communication. The corresponding channels will carry voice, data link or teletype depending on the equipment complement of the aircraft. The manner in which the different classes of communication are assigned to the available satellite capacity requires a detailed system operation study. The previously mentioned Interrogation and Channel Assignment procedure would provide a mechanism for ground control leading to optimized utilization of capacity in response to demand.

Air-to-Ground via Satellite

Let us consider first the possible need for transmission of ranging signals on the air-to-ground link. As mentioned earlier search and rescue operations could be facilitated by the relay of observed navigation signals from a distressed craft. Such an emergency mode could be accommodated on the air-to-ground link via satellite through the temporary cessation of less essential traffic. Under normal operating conditions, the arguments in favor of transmitting ranging signals from aircraft arise from a desire to (1) place the burden of computing position location on the ground instead of on the aircraft, and (2) make air traffic surveillance more nearly independent of aircraft processing equipment.

The benefits of ground computation of position must be weighed in the light of the following factors: possible saturation of satellite capacity and ground facilities with increased air traffic, expanding use of airborne

computers to perform many aircraft functions, and the philosophy of autonomous aircraft navigation without ground intervention.

"Independent" surveillance via satellites is basically a question concerning airborne equipment reliability. The retransmission of ranging signals upon ground request from each aircraft would substantially increase the required satellite capacity. The economics of providing this capacity must be balanced against the cost of maintaining the reliability of aircraft processing equipment at a sufficiently high level. Moreover, the circuits for retransmitting ranging signals are subject to degradation which can cause errors in the aircraft position computed by the ground station. The equipment which in response to an interrogation merely transmits position location as derived by the aircraft may actually be more reliable than a partially independent surveillance system. On balance, we conclude that retransmission of ranging signals from air-to-ground via satellite is an undesirable feature. This implies that the air-to-ground link only carries communication, in particular, multiple accesses to the satellite from different aircraft terminals.

Of the two major contenders -- time division multiple access(TDMA) and frequency division multiple access (FDMA) - TDMA requires in the course of normal operation the equivalent of a range measurement. In order for each user to fall into his assigned time slot at the satellite, he must compensate for the propagation delay on the vehicle-to-satellite path. This can be done in a closed-loop fashion by having the user observe the time of occurrence (relative to a frame sync) of his own signal as it is retransmitted by the satellite. The user can then correct his transmission time so as to fall into the assigned slot.

There are several disadvantages to this form of multiple access. Initial establishment of the link requires a search and acquisition for time synchronization. Where frequent short contacts are the rule as in air traffic control, acquisition and synchronization must be easy and rapid. Note that in TDMA the time slot synchronization for noninterfering multiple access at the satellite is in addition to the normal synchronization procedures involved in the reception of digital data. A second disadvantage of TDMA is that it forces the aircraft to receive the satellite-to-ground signal for the sole purpose of synchronizing the multiple access channels. Since the satellite-to-ground station transmission is intended for the control center, there is no need to provide aircraft with capability of receiving this signal. In fact, it is probably advantageous to employ two different frequency bands for the two satellite links. Since the power budget on the satellite-to-ground link involves a high gain antenna at the ground station, it would be uneconomical from the standpoint of satellite power to make this signal also receivable at the aircraft. This argument along makes TDMA unattractive for this multiple access application.

Under the heading of FDM there are still a number of options regarding the type of signals transmitted by the users and the processing performed at the satellite. To avoid problems of power equalization among users, it is desirable to independently limit each user signal arriving at the satellite. This implies that these signals should be angle modulated. Since it is well also to angle modulate the satellite transmitter, one arrives at a PM/PM signal format where the main carrier is modulated by a composite baseband consisting of independently limited user PM signals.

The air-to-ground communications traffic will consist of the complement of the types already discussed for the ground-to-air link. In response to interrogations from the air traffic control system, the aircraft will send position reports and other pertinent flight plan information. Environmental conditions will be reported for use in meteorological predictions. Aircraft status will be transmitted to cognizant airline monitoring centers. Much of this flow of information can be handled by automatic read-out of aircraft instruments with little or no intervention by flight crews. In addition, there should be provision for two-way voice communication and teletype to any of several ground sites including all of the above plus connection to common carrier networks.

There must be procedures for aircraft to initiate transmissions in the course of normal operations. This can be accomplished by means of a specific air/ground channel over which aircraft can request channel assignments to a desired ground destination. Like Interrogation and Channel Assignment from ground-to-air, this channel would be for low duty cycle utilization and can be time-shared among many users. However, because calls do not originate at a single point, possible conflicts arise when two or more aircraft transmit simultaneous requests. It would be impractical for each aircraft to listen to the satellite-to-ground transmission to establish whether or not the request channel is clear. Therefore, the system must be able to tolerate occasional simultaneous requests or else the ground station must provide an indication on the ground-to-air link when the request channel is busy.

Concluding Remarks

The preceding paragraphs have provided some general orientation on the systems problems associated with a unified ground-air-ground communication/navigation systems. A number of generalizations concerning these systems have been reached on the basis of qualitative arguments. It is clear, however, that the definition of a specific configuration must await a more detailed study into the requirements and procedures of several potential users of the system. Pending the collection of the necessary data, effort has been concentrated on the major new item of the unified system, viz., modulation techniques for combined communication and ranging signals. The following section provides technical discussion and trade-offs in the area.

SECTION 3

COMBINED RANGING AND COMMUNICATION SIGNALS

This section discusses several aspects of combined communication/ranging modulation designs. In particular, the properties of pseudorandom ranging codes are presented. The potential for transmitting data with such codes is discussed and the tradeoffs presented. Then a straightforward method for deriving ranging data from a high data rate communication signal is investigated and analytic expressions for ranging accuracy are derived. The section closes with a discussion of harmonic ranging systems and their communication capability. The tradeoffs associated with these systems are discussed.

Pseudorandom Range Code Systems

A pseudorandom code ranging system operates by transmitting a binary code with special correlation properties. This code is continuously repeated by the transmitter. There are various methods by which the transmission of data may be combined with such a ranging signal. Several such methods will be taken up in subsequent sections. First, however, this section reviews the properties of this form of ranging.

The properties of pseudorandom sequences are well known.* These sequences may be generated readily using shift registers, and have the desirable property that their autocorrelation (when computed over the length of a code period) has the form shown in Fig. 1. The bit duration is T and the code is $k = 2^n-1$ bits in length, where n is an integer.

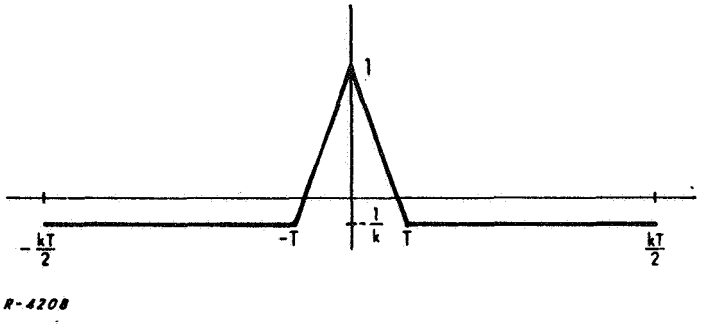
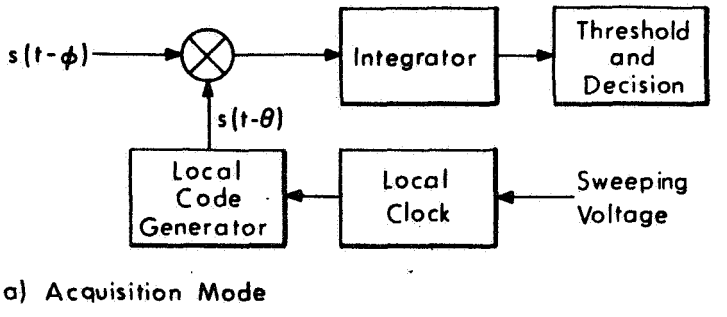


Fig. 1 Autocorrelation of Pseudorandom Ranging Code

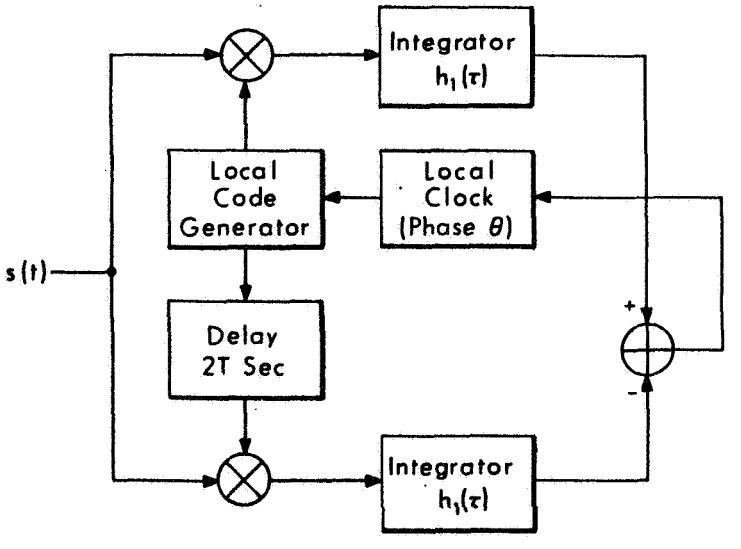
* See, for example, "Digital Communications with Space Application," S.W. Golomb, et al., Prentice-Hall, Englewood Cliffs, N.J., 1964.

In order to preserve generality, the method of modulation used to transmit the pseudorandom sequence will not be specified. Appropriate modulation methods would include phase-reversal keying (PRK) and frequency-shift keying (FSK). Details of receiver operation will differ with different modulation methods. Therefore, receiver operation will be discussed as if all operations took place at baseband. Differences in performance will also occur for different modulation schemes; however, it will be necessary to take some particular results in order to analyze the trade-offs involved.

A ranging receiver operates in two modes; acquisition and tracking. The principle of operation is illustrated by Fig. 2. In the acquisition mode,



a) Acquisition Mode



b) Tracking Mode

R-5410

Fig. 2 Typical (Baseband) Receiver Structure for Pseudorandom Code Ranging

the incoming code is correlated with a locally generated replica of the code. The phase of the local code is swept slowly until the output of the correlation operation exceeds a threshold, indicating the the two codes are approximately in phase. During the acquisition process, two types of malfunction may occur; false lock due to noise, and failure to lock. Analysis of the acquisition process in one form of receiver using PRK modulation gives the result shown in Fig.3. The performance depends on the product $T_S(S/N_o)$ where T_S is the time to sweep through one bit period and S/N_o is the ratio of signal power S to (one-sided) RF noise density. To ensure successful acquisition the total probability of Fig.3 must be constrained. A reasonable constraint (in the absence of any data modulation) is

$$T_s \frac{S}{N_o} \geq 25 \quad (14 \text{ dB}) \tag{1}$$

Since average acquisition time is given by

$$T_{acq} = \frac{1}{2} kTs, \tag{2}$$

Equation (1) may be rewritten as

$$\frac{T_{acq}}{k} \frac{S}{N_o} \geq 12.5 \tag{3}$$

Thus there is a direct trade-off between acquisition time, code length, and signal power-to-noise density ratio. This equation (with the equality) is plotted in Fig. 4.

The tracking mode of operation is shown in Fig. 2b. The output of the summer is an error signal which controls the phase of the local clock. Figure 5 shows the magnitude of this voltage as a function of phase offset in the absence of noise. In Appendix B it is shown that the rms range error of such a system is given by:

$$\Delta r = \frac{cT}{\sqrt{T_{int} (S/N_o)}} \tag{4}$$

where c is the velocity of light and T_{int} is the integration time of the tracking loop. This is inversely proportional to the loop bandwidth. Normally, loop bandwidth cannot be reduced beyond a certain point due to the presence

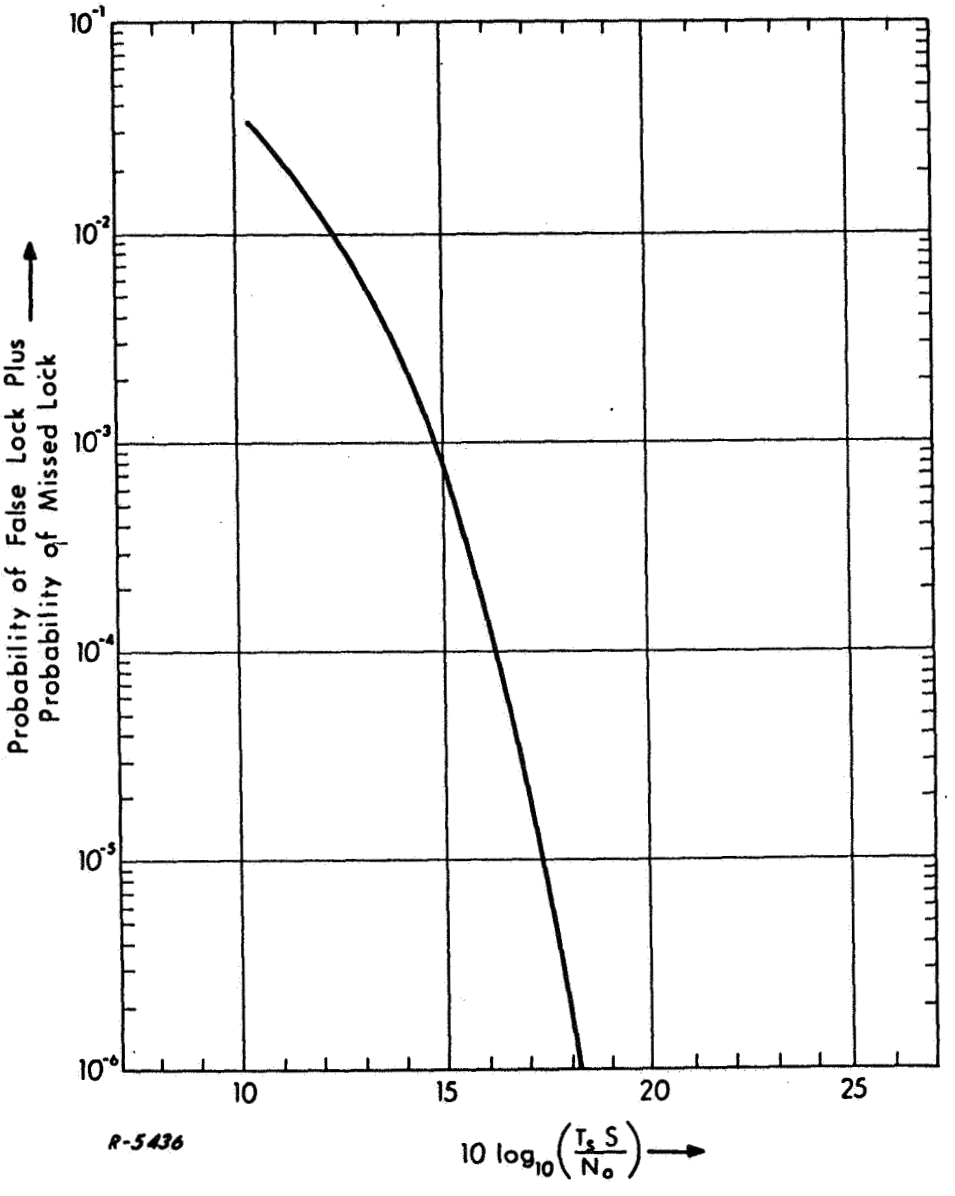


Fig. 3 Acquisition Performance of Pseudorandom Ranging System

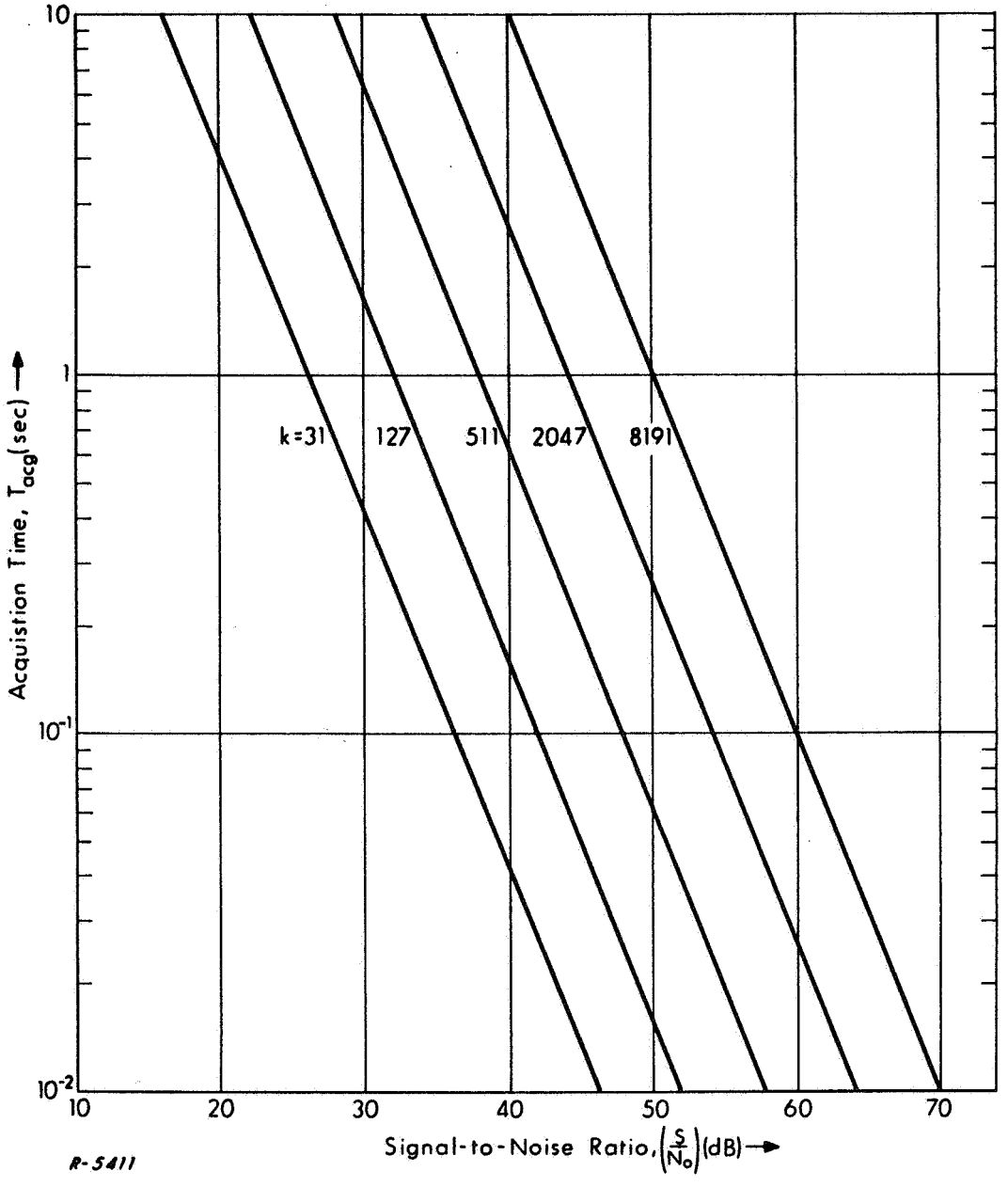


Fig. 4 Bound Involving Acquisition Time, Code Length, and SNR (Equation 3)

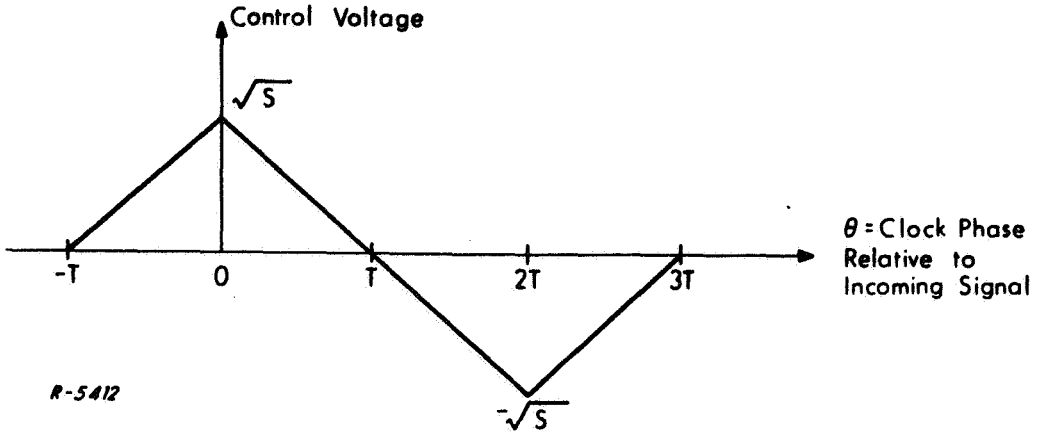


Fig. 5 Control Characteristic in the Tracking Mode

of doppler shifts in the received signal. Therefore, Fig. 6 shows Δr as a function of bit duration and signal-to-noise ratio with T_{int} equal to 10^{-2} sec.

The restrictions on signal-to-noise ratio due to limited satellite power, high path loss, etc. are obvious. Not so obvious are the consequences of short bit duration as a means of achieving accuracy. To begin with, higher bit rates mean increased bandwidths. For example, for PSK modulation the transmitted bandwidth is approximately

$$BW_{PSK} = \frac{3}{T} \tag{5}$$

Second, increased pulse rate aggravates the problem of range ambiguity. The range determination made by a pseudorandom code is inherently ambiguous due to the finite code length.

The unambiguous range is given by

$$r_{max} = c\,k\,T \tag{6}$$

This relationship is shown in Fig. 7. Reducing T will reduce r_{max} unless the code length k is increased. This in turn will (from Eq. 3) increase acquisition time unless the signal-to-noise ratio is increased.

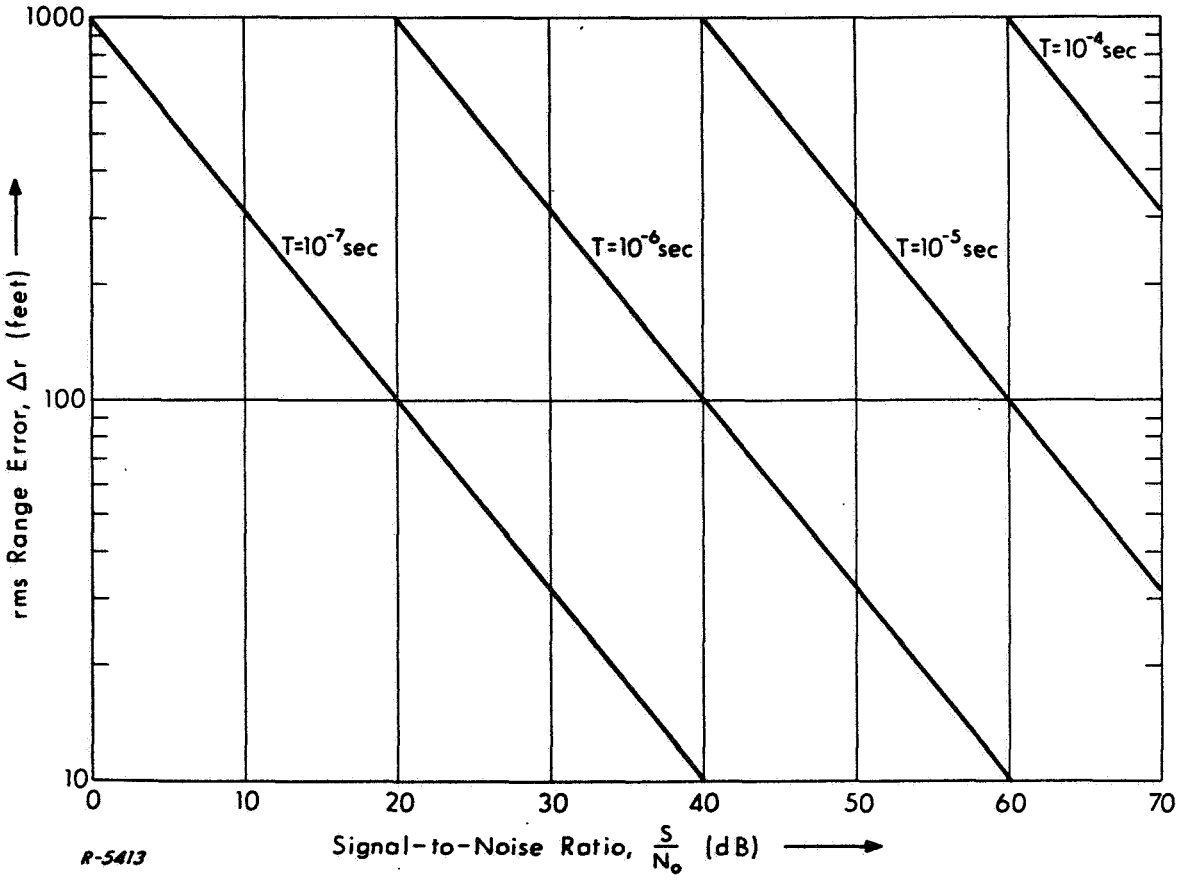


Fig. 6 Ranging Accuracy as a Function of Bit Duration and SNR (Assuming $T_{int} = 10^{-2}$ sec in Eq. 4)

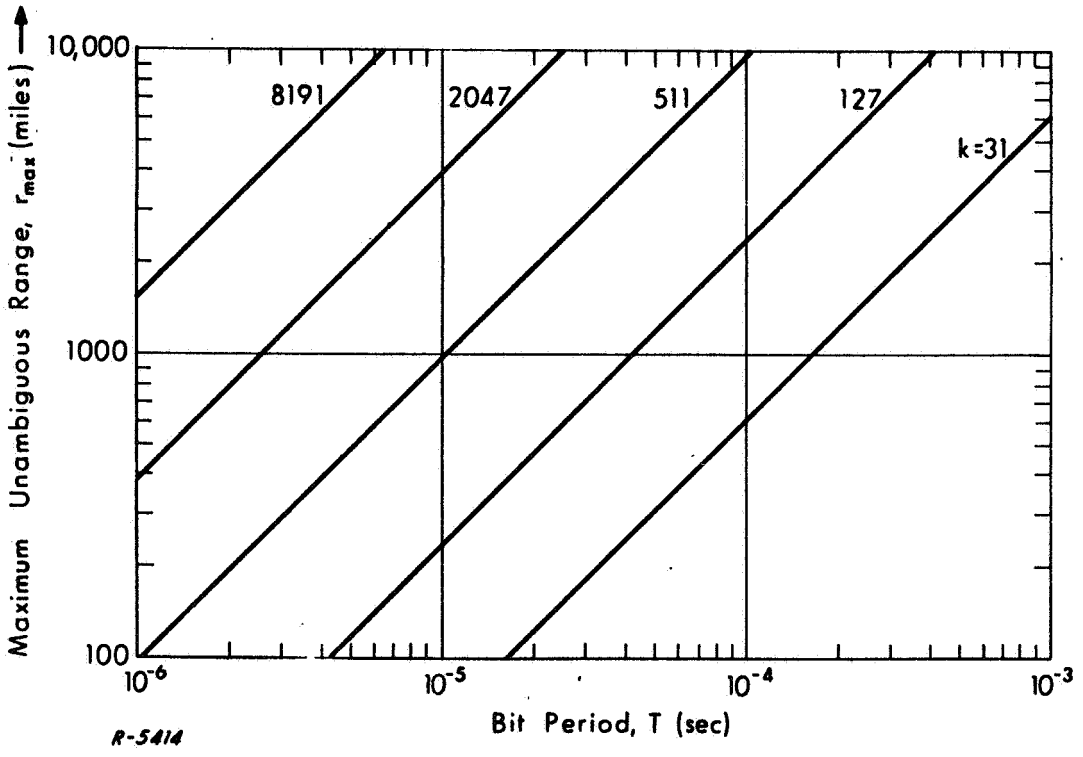


Fig. 7 Relationship Between Maximum Unambiguous Range, Code Length, and Bit Period

Range Code Polarity Reversal

One method of adding data handling capability to a pseudorandom ranging signal is to modulate the polarity of the code. That is, each repetition of the code is sent with either normal or reversed polarity depending on the polarity of the binary data element. Obviously, the bit rate obtained by this process is equal to the repetition rate of the code; therefore, we may relate data rate to the maximum unambiguous range:

$$r_{max} \quad \frac{1}{T_{data}} = c \tag{7}$$

The trade-off between these two quantities is shown graphically in Fig. 8. Clearly the data capacity of this system is limited. If the ranging function

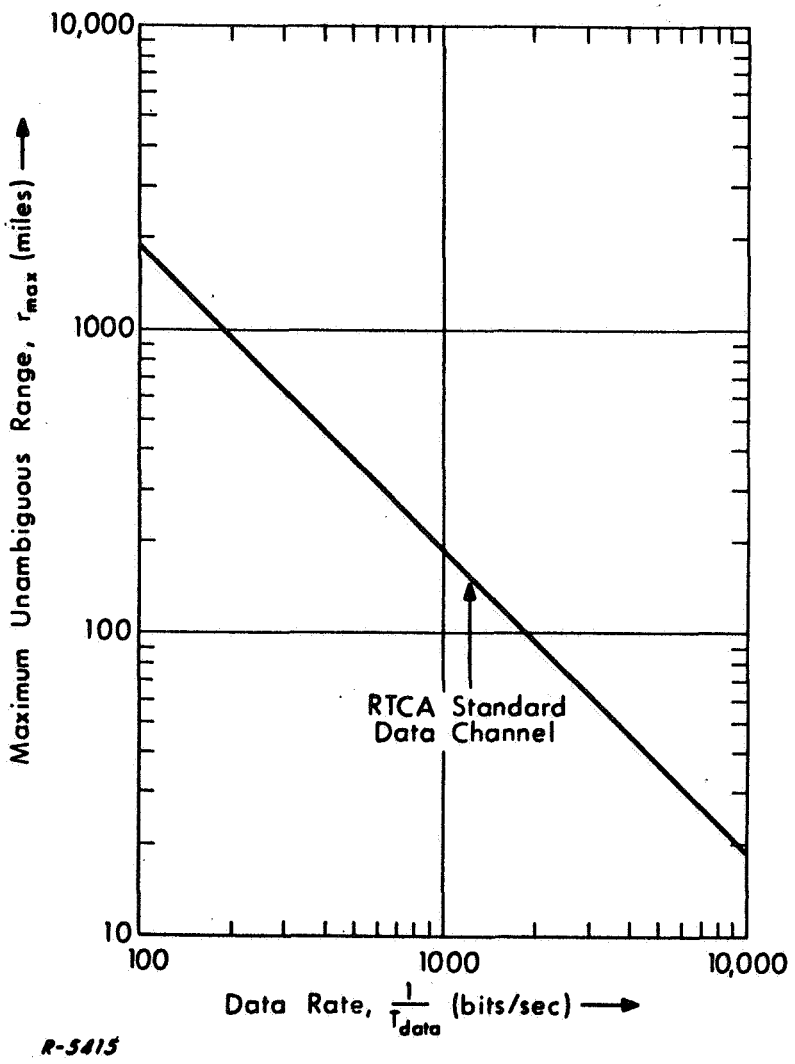


Fig. 8 Trade-off Between Data Rate and Maximum Unambiguous Range

is required to resolve ambiguities over a substantial part of a hemisphere, say an r_{max} of 1000 or 2000 miles, the data capacity is only adequate for special purposes. On the other hand, the ranging function could be used to increase the accuracy of other navigation systems. In that case the other techniques could provide ambiguity resolution down to 100 miles, enabling the pseudorandom ranging system to carry a 1200 bit/sec data channel.

Addition of data to the pseudorandom ranging system in this manner will cause some degradation of ranging performance. Specifically, the code will no longer have the desirable correlation properties of Fig. 1. Figure 9 shows one example of this, the correlation of a 31 bit code with its complement. This shows up primarily as a loss of acquisition performance, since the threshold level must be raised to avoid false lock on the local maxima of the correlation function. Therefore, Eq. (3) cannot be applied to this type of combined link without increasing the constant on the right-hand side.

The error probability of the binary data transmitted by this scheme is related to the ratio E/N_o. The quantity E is the transmitted energy per data bit and N_o is the noise density defined previously. In terms of the quantities defined in the section entitled, "Pseudorandom Range Code Systems," on p.13, this is given by

$$\frac{E}{N_o} = \frac{kTS}{N_o} \tag{8}$$

Figure 10 gives the relationship between E/N_o and error probability for various modulation methods. The modulation method used for the range code bits determines that of the data. Therefore, if the range code is phase reversal keyed and coherently detected, the lowest curve of Fig. 10 applies, since the two data symbols are antipodal (correlation of -1). FSK modulation, on the other hand, is orthogonal (correlation of 0 on each code bit) thus more energy per data bit is required. For PSK Fig. 10 indicates how large the ratio E/N_o must be for acceptable data quality. This in turn sets a requirement on kTS/N_o from Eq. (8).

Ranging from Bit Synchronization

All digital data transmission links carry timing information in one form or another, since timing information is necessary for detecting the data. Several levels of timing may be contained in the transmitted signal and extracted by the receiver. If the system is coherent, the receiver tracks the actual carrier phase. Whether or not the system is coherent, the receiver must derive bit timing in order to properly detect the data. In a high capacity link, the data may be time-multiplexed, requiring the use of a "frame" structure. Frame synchronization patterns at the start of each frame allow

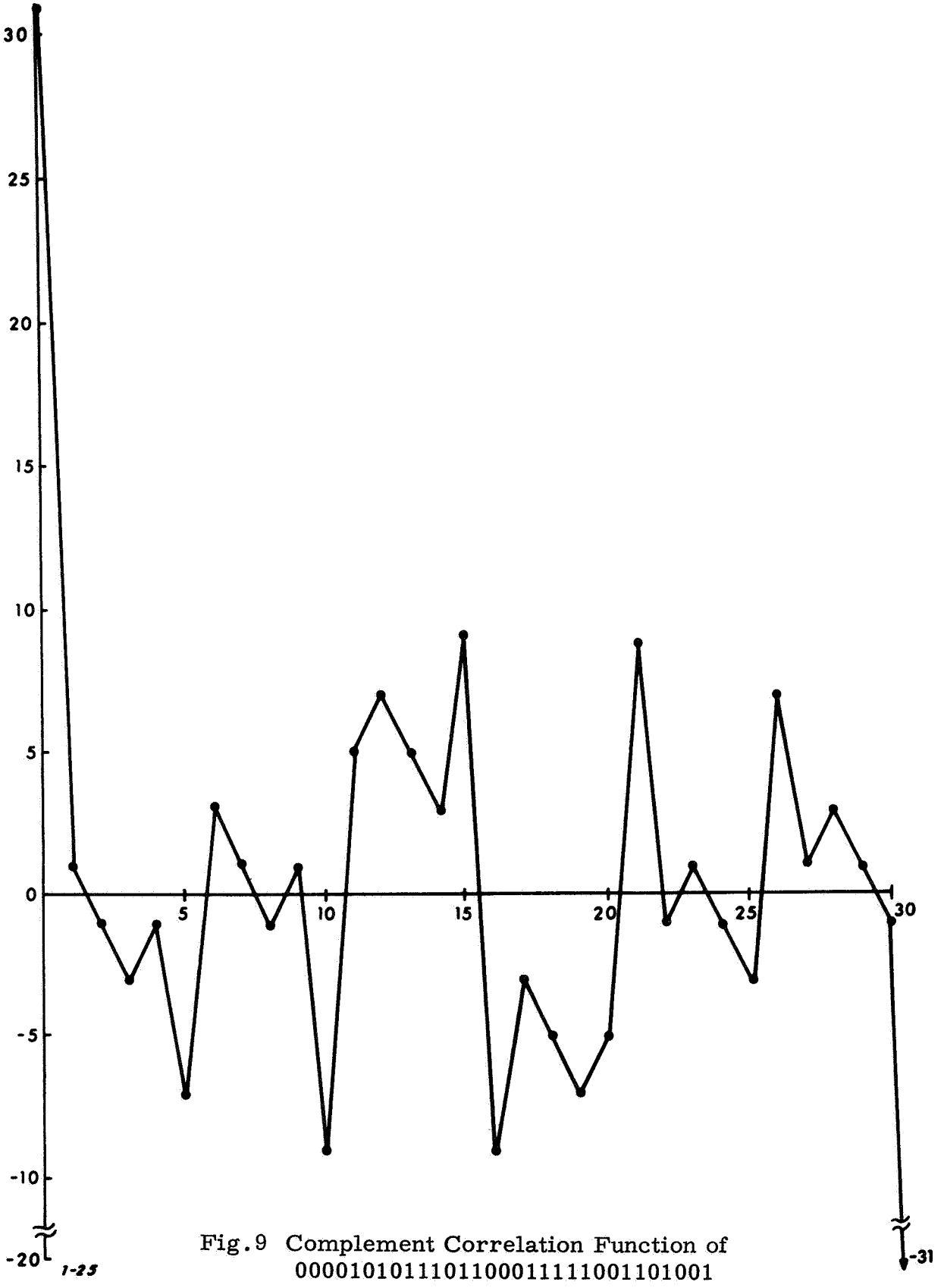


Fig. 9 Complement Correlation Function of
00001010111011000111111001101001

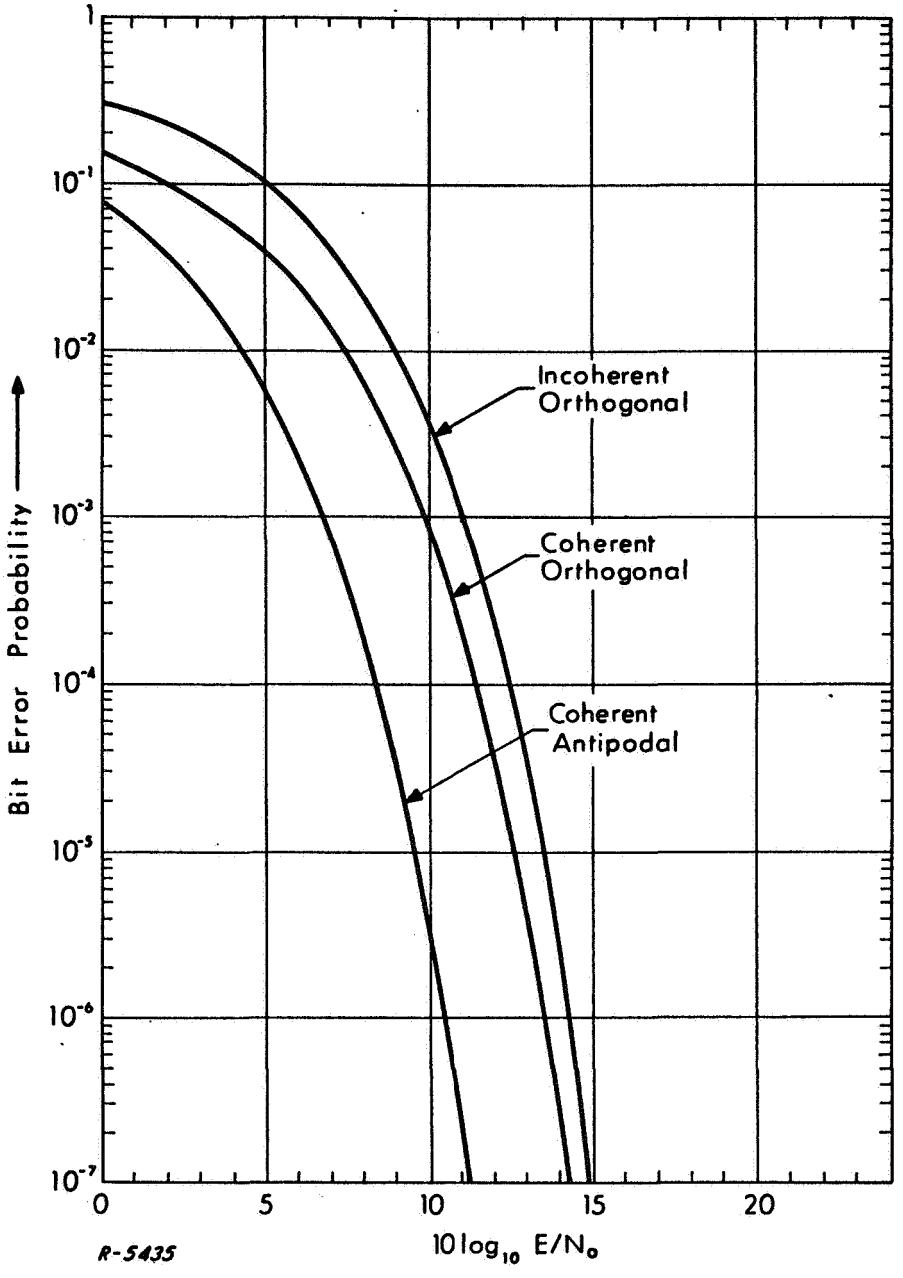


Fig. 10 Error Performance of Various Binary Modulation Methods

the receiver to maintain frame synchronization and de-multiplex the data properly. In a system which is transmitting messages sequentially to many users, the frame sync patterns might be replaced with message addresses. However, if these could be transmitted at uniform intervals, they would contain the same type of timing information.

In view of the above, it is quite possible that a unified ranging/data system could operate using a transmitted signal identical to that which would be used by a data link above. The suggested principle of operation is that the bit timing derived by the receiver could be used for fine ranging while counting bits relative to the frame sync pattern would give the coarse range information. The purpose of this section is to look into the performance obtainable with such a system. Unfortunately, no general treatment of the problem of sync extraction is available. From a practical point of view, many techniques of synchronization are available, as well as various modulation techniques. Therefore, only one example will be analyzed.

Generally speaking, sync extraction requires a non-linear operation on the received signal. This is because the spectrum of a randomly modulated carrier will not contain a discrete line at the bit rate. The system to be analyzed is typical in this respect. Consider a received signal of the form:

$$S(t) = 2A(t) \sin \frac{\pi t}{T} \sin \omega_c t \quad 0 \le t \le T \tag{9}$$

where A(t) is either +A or -A during any interval (0, T). From the standpoint of data detectability, this is an efficient method of transmitting data. Performance of a coherent receiver in the presence of noise is given by the lowest curve of Fig. 10, since the binary signals have a correlation of -1. The structure of a coherent receiver is shown in Fig. 11. The output of the squarer contains periodic components at twice the carrier frequency and twice the symbol frequency, since

$$S^2(t) = A^2 \left[1 - \cos \frac{2\pi t}{T} - \cos 2\omega_c t + \frac{1}{2} \cos 2 (\omega_c - \frac{\pi}{T}) t + \frac{1}{2} \cos 2 (\omega_c + \frac{\pi}{T}) t\right] \tag{10}$$

The receiver tracks both of these components using phase-locked loops. Viterbi* has analyzed the performance of the carrier loop. In Appendix C his analysis is extended to cover the bit sync loop as well. The result obtained is that the rms loop error (converted to a range measurement) is

---

*"Principles of Coherent Communication," A. J. Viterbi, pp. 286-292, McGraw-Hill, New York, 1966.

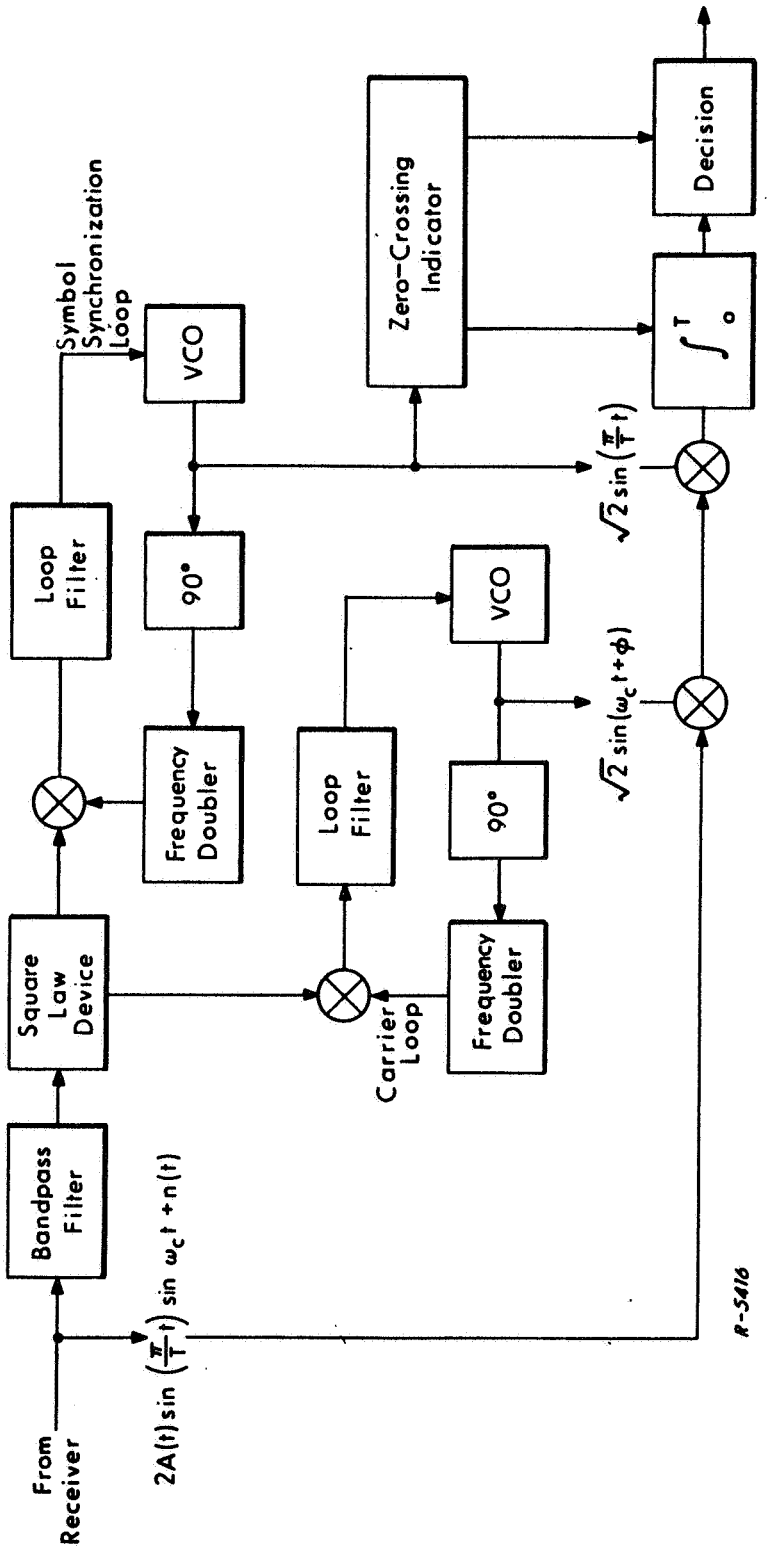


Fig. 11 Receiver Structure Including Bit Sync Extraction

$$\Delta r = \frac{cT}{4\pi}\sqrt{TB_L} \tag{11}$$

where B_L is the bandwidth of the symbol synchronization loop. This result is based on a bit energy-to-noise density ratio (E/N_o) of 10 dB. From Fig. 10, this value will be necessary for high quality data transmission.

Equation (11) is plotted in Fig. 12. This graph makes it possible to determine how high the data rate must be in order to use bit synchronization for ranging. For instance, if desired ranging accuracy is 30 feet and a loop bandwidth of 20 Hz can be used, a data rate of 85 k bits/sec or greater will be necessary. The loop bandwidth which can be used depends on the frequency stability of the source generating the data stream and on the magnitude of the signal dynamics caused by aircraft and satellite motion. The higher bit rates will probably require larger loop bandwidths.

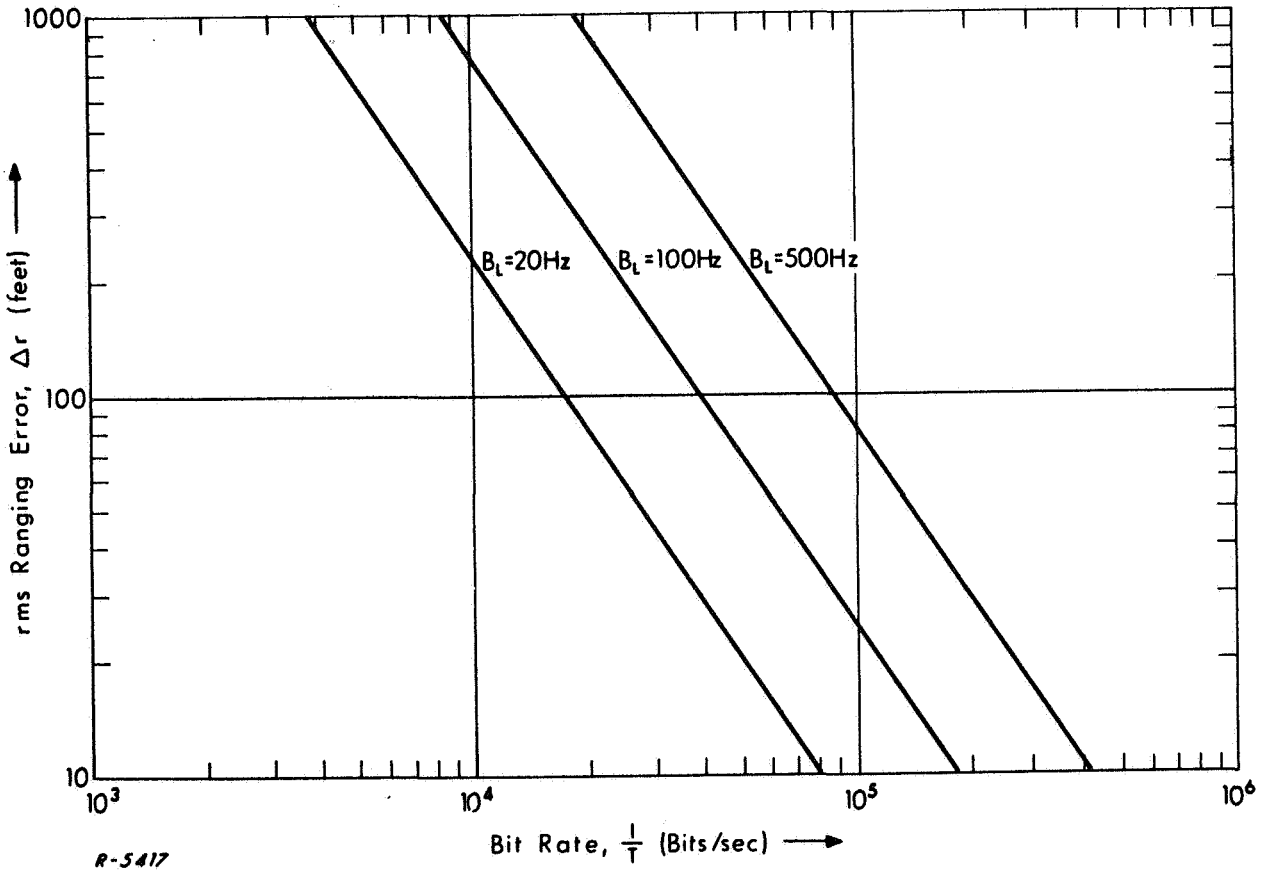


Fig. 12 Ranging Accuracy of Bit Sync System

Combined Harmonic Ranging and Data Transmission

A harmonic ranging system operates by measuring the phases of several harmonically related sinusoids. In a simple case, three frequencies, f_1, f_2 and f_3 of the tones might be in the ratio 64:8:1. By measuring the phase of the f_1 tone, the receiver determines the range precisely, but in a highly ambiguous manner. Measurements of the phases of f_2 and f_3 reduce the ambiguity to a tolerable level, since these frequencies have much larger wavelengths. To continue this example, the actual transmitted signal might consist of a low-deviation PM modulated carrier:

$$S(t) = \cos (\omega_c t + m_1(t)) \tag{12}$$

The modulation $m_1(t)$ could consist of a subcarrier (frequency f_1) which is AM modulated by a subcarrier (frequency f_2) which is in turn AM modulated by a sinusoid at frequency f_3. That is,

$$m_1(t) = \delta_1 \left[1 + \delta_2 (1 + \delta_3 \cos \omega_3 t) \cos \omega_2 t\right] \cos \omega_1 t \tag{13}$$

The modulation indices δ_1, δ_2, and δ_3 are typically on the order of 1/2. The spectrum of the modulation $m_1(t)$ consists of seven spectral lines, as shown in Fig. 13a. The transmitted spectrum is more complex as can be seen from the alternate form for the transmitted signal:

$$S(t) = \sum_{n_o} \cdots \sum_{n_6} J_{n_o}(\delta_1) J_{n_1}\left(\frac{\delta_1\delta_2}{2}\right) J_{n_2}\left(\frac{\delta_1\delta_2}{2}\right) J_{n_3}\left(\frac{\delta_1\delta_2\delta_3}{4}\right)$$

$$\cdot J_{n_4}\left(\frac{\delta_1\delta_2\delta_3}{4}\right) J_{n_5}\left(\frac{\delta_1\delta_2\delta_3}{4}\right) J_{n_6}\left(\frac{\delta_1\delta_2\delta_3}{4}\right) \cos\Big[\omega_c + n_o\omega_1 + n_1(\omega_1 + \omega_2)$$

$$+ n_2(\omega_1 - \omega_2) + n_3(\omega_1 + \omega_2 + \omega_3) + n_4(\omega_1 + \omega_2 - \omega_3)$$

$$+ n_5(\omega_1 - \omega_2 + \omega_3) + n_6(\omega_1 - \omega_2 - \omega_3)\Big] t \tag{14}$$

where $n_o, \ldots n_6 = 0, \pm 1, \pm 2, \ldots$

This signal is not ideally suited for data transmission because there are a large number of intermodulation product components in the vicinity of the carrier. An alternate form of this technique eliminates this difficulty.

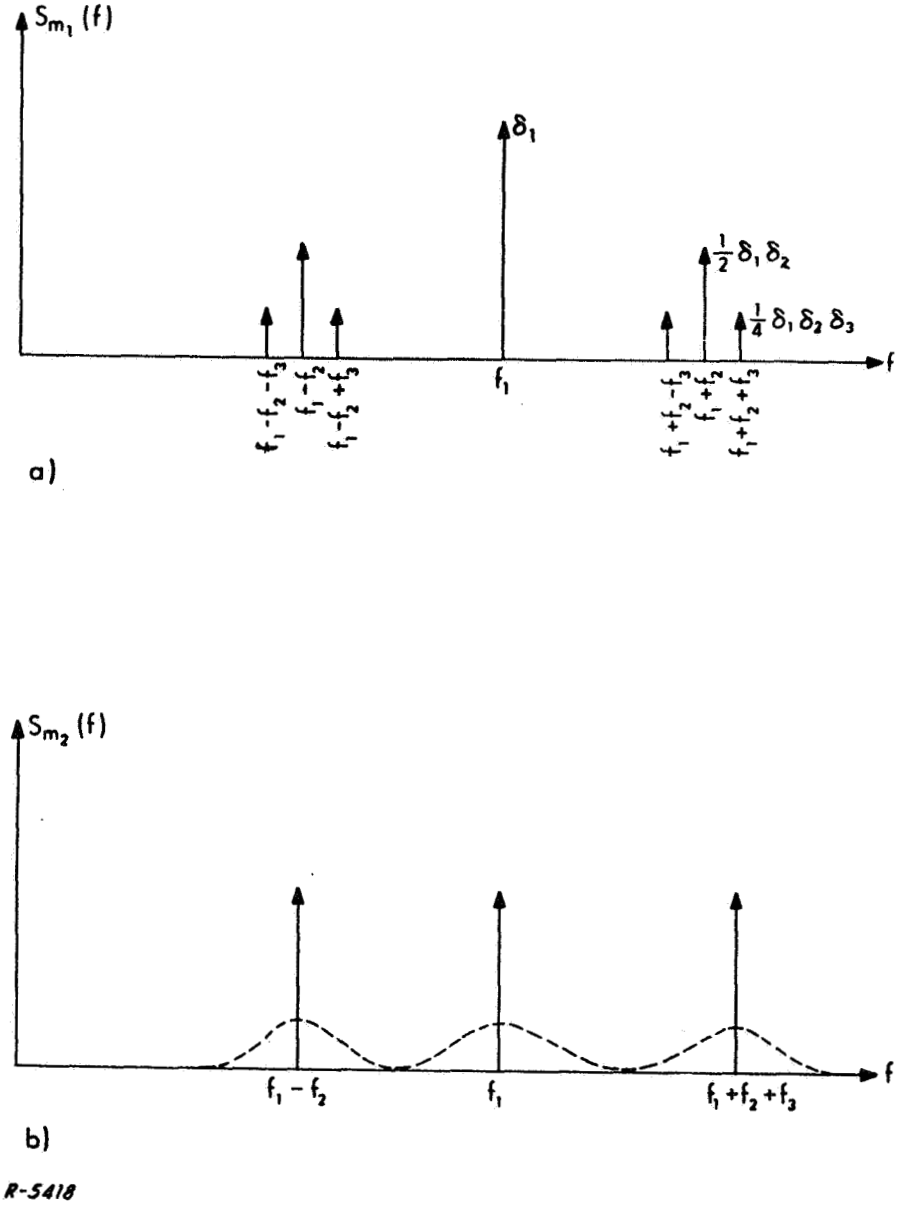


Fig. 13 Spectra of Modulating Waveforms

Specifically, the components at ω_2 and ω_3 could be used to phase modulate the subcarrier at ω_1.* In this case, the spectral region around the carrier is relatively free of intermodulation products and the carrier may be phase modulated by an additional signal which provides communication capability. This configuration is described by

$$m_1'(t) = \delta_1 \sin (\omega_1 t + \delta_o \sin \omega_2 t + \delta_o \sin \omega_3 t) + \phi_{data}(t)$$

For the case where the carrier is modulated directly, $\phi_{data}(t) = \delta_d a(t)$, $a(t) = \pm 1$ is a binary data bit stream. It may be desirable to channelize the data on a number of subcarriers. Then, $\phi_{data}(t)$ would have the form $\phi_{data}(t) = \sum_i a_i(t)\delta_d \sin \omega_i t$, $i = 1, 2, \ldots n$ and n is the number of channels. A variation on this approach consists of using the ω_i subcarriers for ranging purposes and deleting the first term in $m_1'(t)$ altogether. This technique is advantageous in high data rate systems where data bit error rate constitutes the predominant constraint on signal power requirements. To illustrate this technique, consider a modulating signal consisting of three sinusoids:

$$m_2(t) = \cos \omega_1 t + \cos (\omega_1 - \omega_2) t + \cos (\omega_1 + \omega_2 + \omega_3) t \qquad (15)$$

The spectrum of this signal is considerably simpler as can be seen from Fig. 13b. As a consequence, the tones can be modulated by data with a minimum of interference between the two forms of transmission. This signal contains all of the range information available from $m_1(t)$. The three tone frequencies f_1, f_2, and f_3 can be recreated by appropriately mixing the three tones extracted from $m_2(t)$.

With phase-reversal keying, the spectrum of a modulated tone is shown in Fig. 14. The spacing between tones in Fig. 13b is about f_2 Hz, and the width required by the spectrum of Fig. 14 is about $4/T$ (where $1/T$ is the data rate on a single tone). Therefore, each tone could carry a data rate of approximately $f_2/4$ bits/sec. Then the total data rate (using all three tones as separate data subcarriers) is $3f_2/4$.

The system described above offers great flexibility of power division between range tone and data. Consider for example a single tone and the associated continuous (data) spectrum. The ratio between the subcarrier (tone) power S_{sc} and the power in the continuous (data) spectrum S_d may be adjusted to any desired value to give the desired balance between ranging accuracy and data quality.

---

*Variations on this approach are presently under investigation for NASA ERC Contract No. NAS12-567.

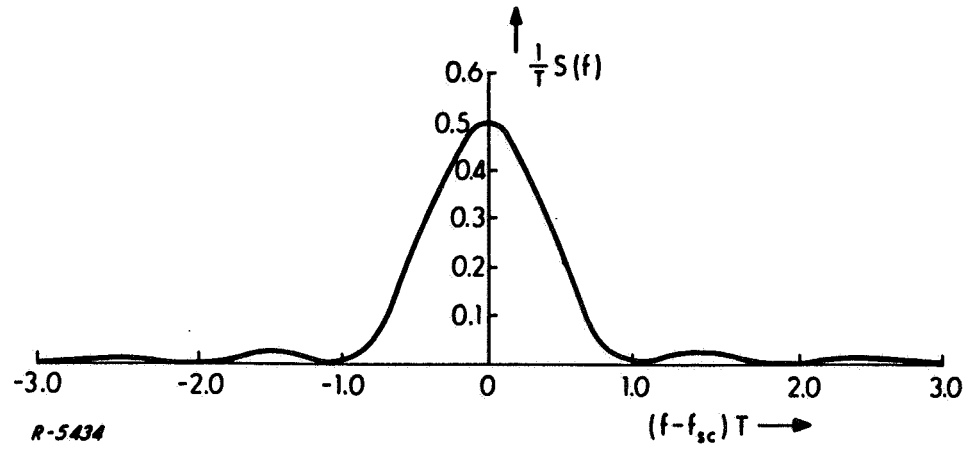


Fig. 14 Spectrum of Data Modulation

One drawback of this system is that the data modulation degrades the performance of the range tone. Assuming that the RF signal has been coherently detected, the additive noise density in the neighborhood of the subcarrier is $N_o/2$ (double-sided). However, as far as range-tone tracking is concerned, the power density of the data signal must also be considered as additive noise. The spectrum of the modulated data (assuming phase-reversal keying) is

$$S_{data}(f) = \frac{S_d T}{2}\left[\frac{\sin \pi (f + f_{sc}) T}{\pi (f + f_{sc}) T}\right]^2 + \frac{S_d T}{2}\left[\frac{\sin \pi (f - f_{sc}) T}{\pi (f + f_{sc}) T}\right]^2 \qquad (16)$$

In the neighborhood of the subcarrier this reduces to a (double-sided) density $S_d T/2$. Therefore, the density of the total disturbance is $(N_o + S_d T)/2$ (double-sided). Then the mean-square phase error* in tracking this subcarrier can be approximated by:

$$\sigma^2_{\phi_{sc}} = \frac{B_L}{S_{sc}}(N_o + S_d T) \qquad (17)$$

* This well-known result for a PLL assumes Gaussian noise which is flat over the bandwidth of the loop. In this case the noise is not Gaussian due to the modulation component. Hence this result is only approximate.

Thus the presence of modulation has increased the mean-square ranging error by the factor

$$\frac{N_o + S_d T}{N_o} \tag{18}$$

The factor of Eq. (18) will generally be appreciable. If an error probability of 10^{-3} is maintained, the ratio E/N_o for the modulated data must be 7 dB or greater. Then

$$\frac{S_d T}{N_o} = 5 \tag{19}$$

so that,

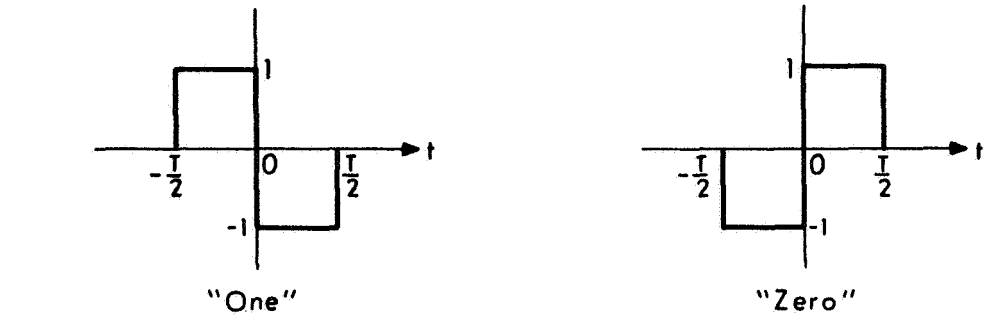
$$\frac{N_o + S_d T}{N_o} = 6 \tag{20}$$

In other words, the data has increased rms ranging error by a factor of $\sqrt{6}$. This degradation in performance can, of course, be offset by increasing ranging power the required amount. An alternate approach is to modify the data spectrum by using a split-phase format. In split-phase signaling, the modulating waveform changes polarity in the middle of each bit, as shown in Fig. 15. The resulting spectrum (Fig. 15b) is zero at the subcarrier; therefore, it will not interfere with ranging. The drawback of this technique is the increased spectral occupancy of the split-phase modulation. Since the data bandwidth is approximately doubled in the split-phase case, the data capacity corresponding to that cited above is reduced from $3f_2/4$ to $3f_2/8$ while ranging accuracy is increased by $\sqrt{6}$.

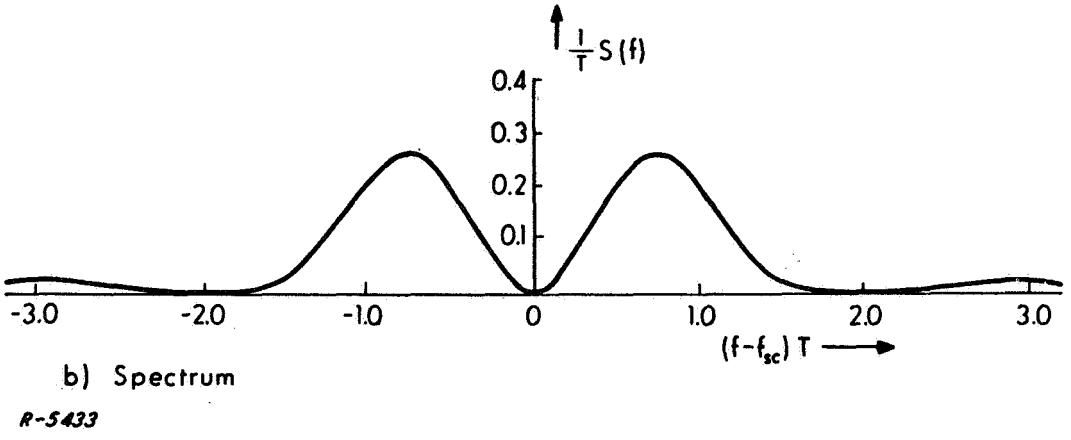
Summary

This section has discussed the properties of several combined ranging and communication modulation techniques. Pseudorandom ranging techniques were discussed first. The acquisition performance and ranging performance were presented. Several tradeoffs involving code length, SNR, acquisition, ranging accuracy, bit period and maximum unambiguous range were derived. It was shown that acquisition time increases with code length, ranging error increases with increased bit length and that if code length and/or bit period are reduced, then, maximum unambiguous range is reduced.

The utility of pseudorandom sequences is discussed in the section entitled, "Range Code Polarity Reversal," p. 21. There it is shown that the data



a) Signalling Waveforms



b) Spectrum

R-5433

Fig. 15 Split-Phase Waveforms and Spectrum

rate is severely limited by maximum unambiguous range requirements and that the presence of data degrades acquisition performance. Bit error probability for various modulation schemes as a function of SNR is presented.

The section entitled, "Ranging from Bit Synchronization," p. 22, derives the ranging performance derivable from a typical high data rate communication signal. A tradeoff of ranging accuracy vs bit rate with tracking bandwidth as a parameter is derived.

The section entitled "Combined Harmonic Ranging and Data Transmission," p. 28, describes a combined communication-ranging technique which consists of a number of sinusoidal subcarriers low deviation modulated by digital data. The interaction between data and ranging performance is discussed. In particular it is shown that the data appears as noise in the ranging extraction process if ordinary data modulation techniques are employed. The degradation in ranging accuracy is derived. Finally, a modified data modulation format is presented which eliminates this difficulty at the expense of reduced data rate.

SECTION 4

NAVIGATION USING TIME REFERENCE SYSTEMS

This section is concerned with the characterization of errors in position determination associated with such techniques as LORAN-C, OMEGA, and DME/DME. Particular emphasis is placed on evaluating the utility of additional one-way passive ranging measurements obtained with the aid of an on-board stable timing reference.

It is shown in the section entitled, "Utilization of Surfaces in Trajectory Situations," p. 37, that several techniques yielding three independent measurements may be employed to determine the position of a vehicle. Consideration is given, specifically, to range-only, range-sum and range-difference methods for the purposes of illustration. Of course, other techniques are very frequently employed such as angle-only and range-angle measurements.

In the discussion entitled, "The Least-Squares Data Reduction Technique," p. 45, the method of least-squares data reduction is derived and presented. This technique provides a generalized approach for combining data obtained from several different types of measurements. The least-squares method also provides a means for optimum position location in the over-determined case. A system of equations is said to be over-determined when there are more independent equations than unknown variables.

Examples are presented in the section entitled, "Investigation of GDOP Phenomena Using the Least-Squares Data Reduction Technique," p. 54, in which the least-squares method is used to determine errors in position location as a function of position, i.e., GDOP errors. Several position determination methods are hypothesized and investigated. Particular emphasis is placed on the improvements in position location performance obtainable if one-way ranging is accomplished with the aid of an on-board stable time reference system. Specific consideration is also given to the use of such techniques and systems for altimetry.

Background

Vehicle position is given by the point of intersection of three or more surfaces. Spherical surfaces are obtained from monostatic or one-way range measurements, ellipsoids are obtained from range-sum measurements, and hyperboloids are obtained from range difference measurements. Examples of such surfaces are shown in Fig. 16 for the case of position determination in a 2 dimensional plane.

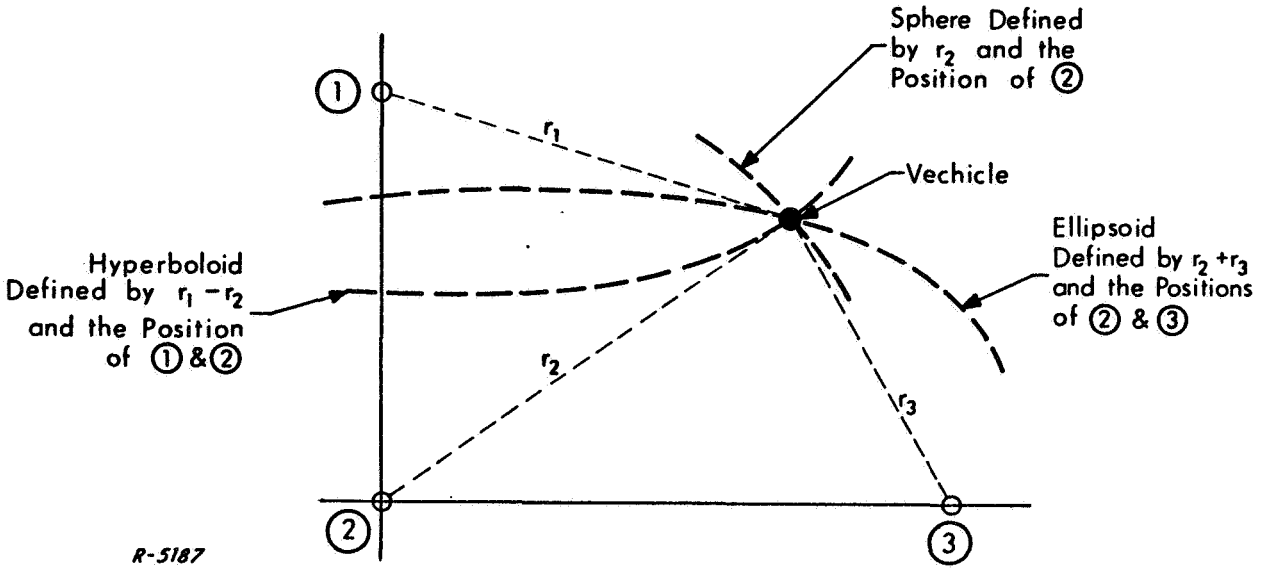


Fig. 16 Intersection of Surfaces Defines Position

Specifically, the spherical surface (shown as a circular line segment in Fig. 16) is defined by a measurement of r_2. The range r_2 may be determined by a round trip delay measurement of the propagation time between the vehicle and station No. 2. Alternatively, r_2 may be determined by a one-way measurement of the propagation delay. This may be accomplished if the time of transmission from station No. 2 is known and if a suitable timing reference is available of the vehicle. In a similar way, spherical surfaces defined by r_1 and/or r_3 may be generated.

Hyperbolic surfaces constitute the loci of all points with a given range difference to a pair of stations. Thus, the station locations are the foci of the hyperboloids. Figure 16 shows a hyperbola defined by the difference $r_1 - r_2$, and the locations of stations No. 1 and No. 2. With configurations shown in Fig. 16 only one more independent hyperbola may be generated, either that defined by $r_1 - r_3$ or $r_2 - r_3$. Range differences are determined, in practice, by measuring the difference in the time of arrival at the vehicle of events transmitted simultaneously from the stations. The events typically consist of the zero crossings of a sine wave or the centers of received pulses, etc.

Note that to obtain three independent surfaces using the range difference technique, 4 stations are required. On the other hand, straightforward measurement of range to the stations yields one independent surface for each station. Note also that the range difference method is particularly suited for passive position determination on-board the vehicle.

The third type of surface illustrated in Fig. 16 is ellipsoidal. Ellipsoids are the loci of all points with a given range-sum to a pair of stations. The range-sum method is not particularly suited to position determination on-board the vehicle. Such techniques are typically employed in bistatic and multistatic radar configurations in which the time of arrival of a transmission from station No. 2, say, is monitored at station No. 1 and No. 3. Thus, $r_1 + r_2$ and $r_2 + r_3$ are determined. Such configurations usually provide a monostatic measurement of r_2 so that r_1 and r_3 may be derived in the data reduction process. In this case, the ellipsoidal surfaces are transformed into spherical surfaces.

As noted above, vehicle position is determined by the point of intersection of at least three surfaces. In general, the surfaces may be spheres, ellipsoids, hyperboloids or any combination of the three. There are certain advantages made possible by an appropriate selection of the surfaces to be used. The optimum combination is found from consideration of several factors including

1) constraints imposed by system configuration;

2) the geometric dilution of precision (GDOP);

3) reduction of bias errors; and,

4) reliability of tracking.

Utilization of Surfaces in Trajectory Situations

It is desirable to obtain the solution of the target position in a manner which minimizes the error in position. The geometric factors play a leading role in reducing the error since any error in measurement can be magnified by GDOP. The basic consideration in obtaining an appropriate geometry is the selection of an appropriate set of ground stations since the geometry depends on the position of the ground stations relative to the target. It might be presumed that for a given set of ground stations there is an advantage in the use of one surface type rather than another. Such an advantage can exist in the case of bias errors. However, this is not the case with geometric factors. A major role of the surfaces is their usefulness in prescribing the necessary ground station locations for a given accuracy requirement. These points are discussed in the following sections which give a description of the uncertainties in the locations of the various surfaces, and derivations of expressions for the target position error which are independent of the choice of surfaces.

Factors Affecting GDOP

GDOP arises in the arithmetic operations used to determine the point of intersection (the target location) of the surfaces. There are two basic

factors which determine the precision with which the intersection can be located, namely, the uncertainty of measured parameters which define the surfaces, and the angle of intersection of the surfaces at the target location. In the case of three orthogonal surfaces the region of uncertainty surrounding the point of intersection is described by a parallelepiped whose dimensions are equal to the uncertainties in the locations of the surfaces. For nonorthogonal surfaces the region of uncertainty is very much greater and grows without bound for the case of nearly parallel surfaces. An appropriate selection of surfaces to be used in the solution for target position is the set of surfaces which are most nearly orthogonal and whose locations are most certain in the region of the target.

The uncertainty in the location of a surface is a function of position on the surface. It may be examined by considering the two dimensional representation in Fig. 17. In the case of an ellipse whose foci lie on the x axis at $x = \pm c$, the uncertainty is greater in the y direction than in the x direction in the region between the foci and near the x axis. At long range, the uncertainties in x and y are essentially equal. In the case of the hyperbola the uncertainty is greatest at distances far from the origin and in those cases where the asymptotes to the curve make a small angle with the x axis.

This may be seen for the case of the ellipse given by

$$\frac{x^2}{a^2} + \frac{y^2}{a^2 - c^2} = 1 \qquad (21)$$

where the range sum $R_s = 2a$. Δy as a function of the error in range sum, ΔR_s is to be found near $x = 0$. Equation (21) then becomes $y^2 = a^2 - c^2$. Differentiating gives $dy = a(a^2 - c^2)^{-1/2} da$ from which

$$\Delta y = \frac{R_s \Delta R_s}{2\sqrt{R_s^2 - D^2}} \qquad (22)$$

where $D = 2c$ is the separation between the two stations. The uncertainty in the x direction, Δx, is found near $y = 0$ by setting $y = 0$ in Eq. (22). Then $x^2 = a^2$ and $\Delta x = \Delta R_s/2$, which shows that Δx is no worse than $\Delta R_s/2$ while Δy near the origin may be much larger. For large values of loop range $\Delta y \approx \Delta R_s/2$ as it should be since the ellipse takes on a form similar to a circle.

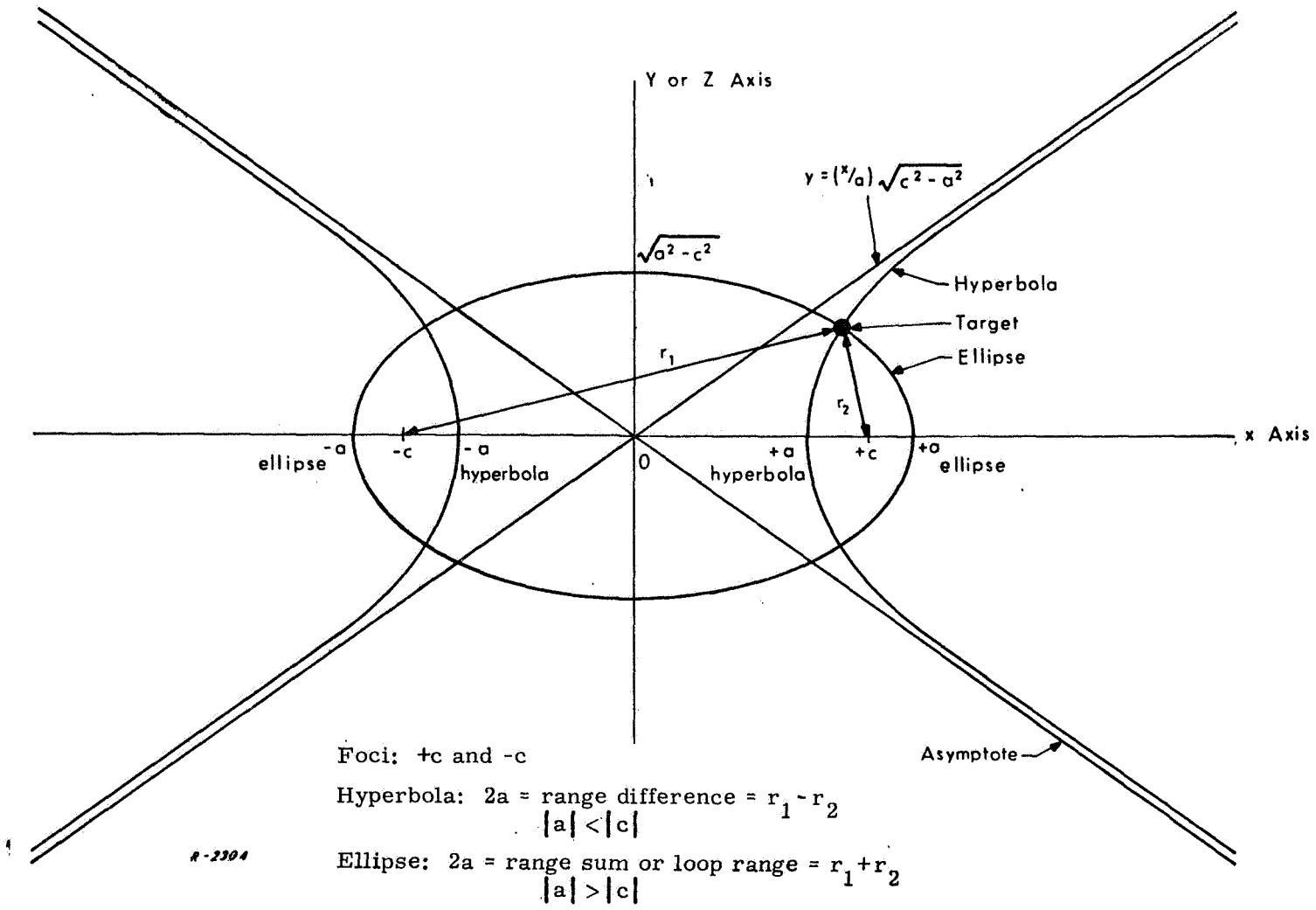


Fig. 17 Geometry of Hyperbola and Ellipse

The uncertainty in the hyperbola of difference range $R_d = 2a$,

$$\frac{x^2}{a^2} - \frac{y^2}{c^2 - a^2} = 1 \tag{23}$$

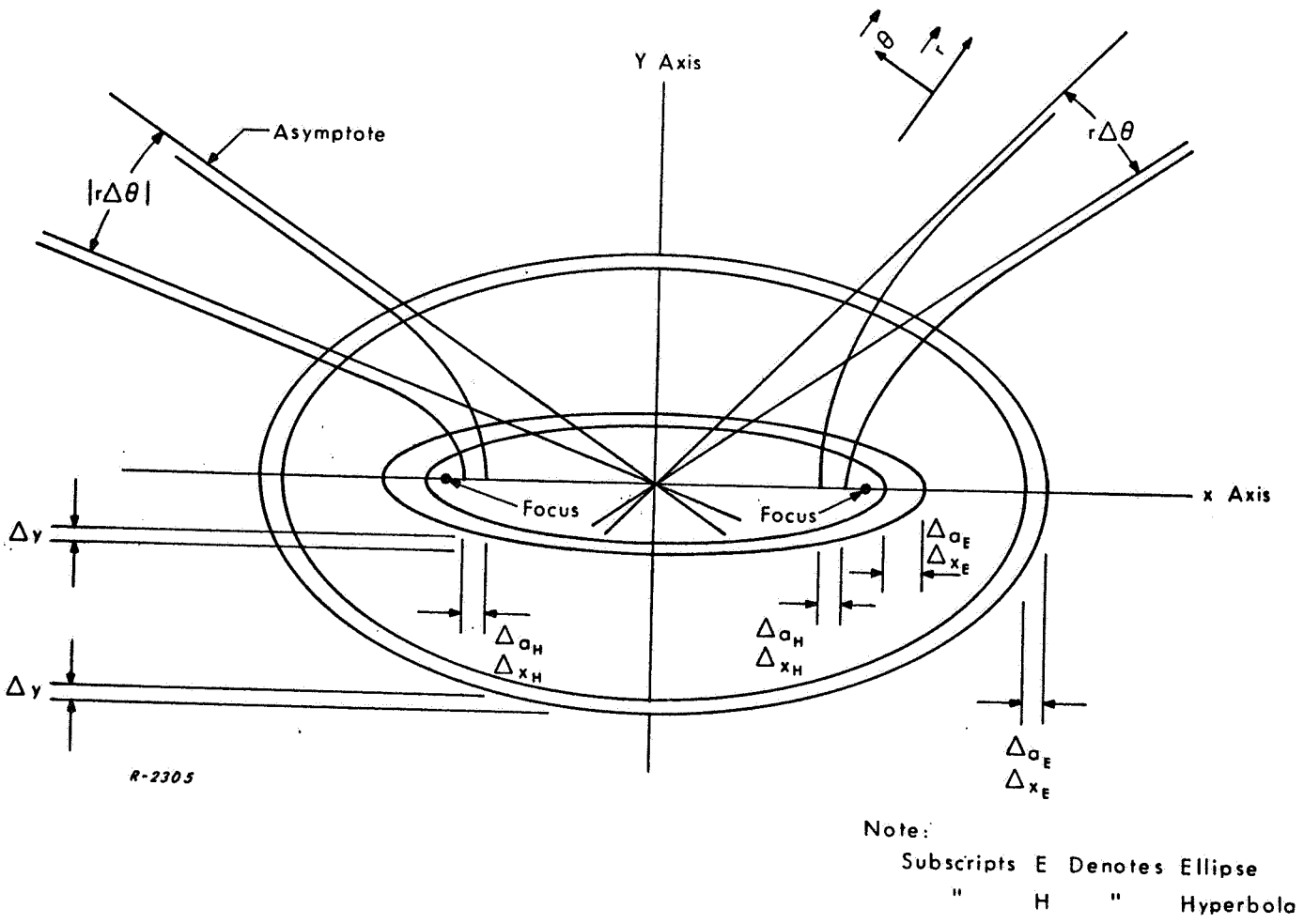
may be treated by an approximation utilizing the asymptotes to the curve and polar coordinates. The asymptote in the first quadrant is given by $y = (x/a)\sqrt{c^2 - a^2}$. The angle θ which the asymptote makes with the x axis is given by $\theta = \tan^{-1}\sqrt{c^2/a^2 - 1}$. The differential change in position in the $\vec{\theta}$ direction is given by $r\,d\theta$ where $\vec{r}$ is the radius vector from the origin. The uncertainty of any point on the asymptote as a function of the uncertainty in range difference ΔR_d is given by $|r\,d\theta| = r\,da(c^2 - a^2)^{-1/2}$ or equivalently by

$$|r\Delta\theta| = \frac{r\Delta R_d}{\sqrt{D^2 - R_d^2}} \tag{24}$$

This approximation is quite accurate for large values of r. The uncertainty given by Eq. (24) is larger than the actual uncertainty in the location of the hyperbola, as may be verified by constructing a second hyperbola corresponding to $(a + \Delta a)$; the displacement in the asymptote is greater than the displacement of the nearest point on the curve. The approximation is of no value near the origin since the displacement of the curve is in the x direction rather than the $\vec{\theta}$ direction. There the displacement may be determined by setting $y = 0$ in Eq. (23) which gives $\Delta x = \Delta R_d/2$.

The uncertainty in the locations of the ellipsoid and hyperboloid may be summarized as follows: For the ellipsoid the minimum uncertainty is equal to $\Delta R_s/2$ and is found at large distances from the x axis as well as right on the axis for $|x| > |c|$. The uncertainty increases without bound in the vicinity of the origin. For the hyperboloid, the minimum uncertainty occurs when the target is located in the region between the foci and near the origin. The uncertainty is then approximately $\Delta R_d/2$. The uncertainty increases without bound for large values of r and/or small θ. This is shown in Fig. 18. In simple terms, the ellipsoid may be used to advantage to solve for target positions which are far from the origin, while the hyperboloid is used for target positions located near the plane passing through the origin and normal to the x axis. At long range the uncertainty in the location of the hyperboloid is greater than that of the ellipsoid.

Two criteria have been mentioned for the selection of an appropriate set of surfaces to minimize the errors associated with GDOP, namely, the set of surfaces which are most nearly orthogonal and whose locations are the



Hyperboloid (one quadrant only) and Ellipsoid; Two Dimensional View

Fig. 18 Uncertainty in Location of Surface

most certain in the region of the target. These two criteria depend on only one set of parameters, the locations of the ground stations relative to the target. This is the case since one may be derived from the other; a knowledge of the surfaces may be used to give the locations of the stations and the ranges from target to each station, or a knowledge of the station locations and the ranges or appropriate combinations of range sums, difference ranges and monostatic ranges may be used to give the surfaces. Insofar as GDOP is concerned, only the locations of the stations relative to the target are important, the surfaces being only a convenience.

Equivalence of Solutions in the Just-Determined Case

As a demonstration of the equivalence of solutions of target position obtained from different sets of surfaces, consider the following examples of a solution in two dimensions between a circle and an ellipse, and a circle and a hyperbola. The equations for the circle and ellipse shown in Fig. 19 are given by

$$x^2 + y^2 = r_1^2 \tag{25a}$$

and

$$\frac{(x - D/2)^2}{(R_s/2)^2} + \frac{y^2}{(R_s/2)^2 - (D/2)^2} = 1 \tag{25b}$$

with solution

$$x = (D^2 - R_s^2 \pm 2r_1 R_s)/2D \tag{26a}$$

$$y = \left[r_1^2 - (D^2 - R_s^2 \pm 2r_1 R_s)^2/4D^2\right]^{1/2} \tag{26b}$$

The equations for the circle and hyperbola shown in Fig. 19 are given by

$$x^2 + y^2 = r_1^2 \tag{27a}$$

and

$$\frac{(x - D/2)^2}{(R_d/2)^2} - \frac{y^2}{(D/2)^2 - (R_d/2)^2} = 1 \tag{27b}$$

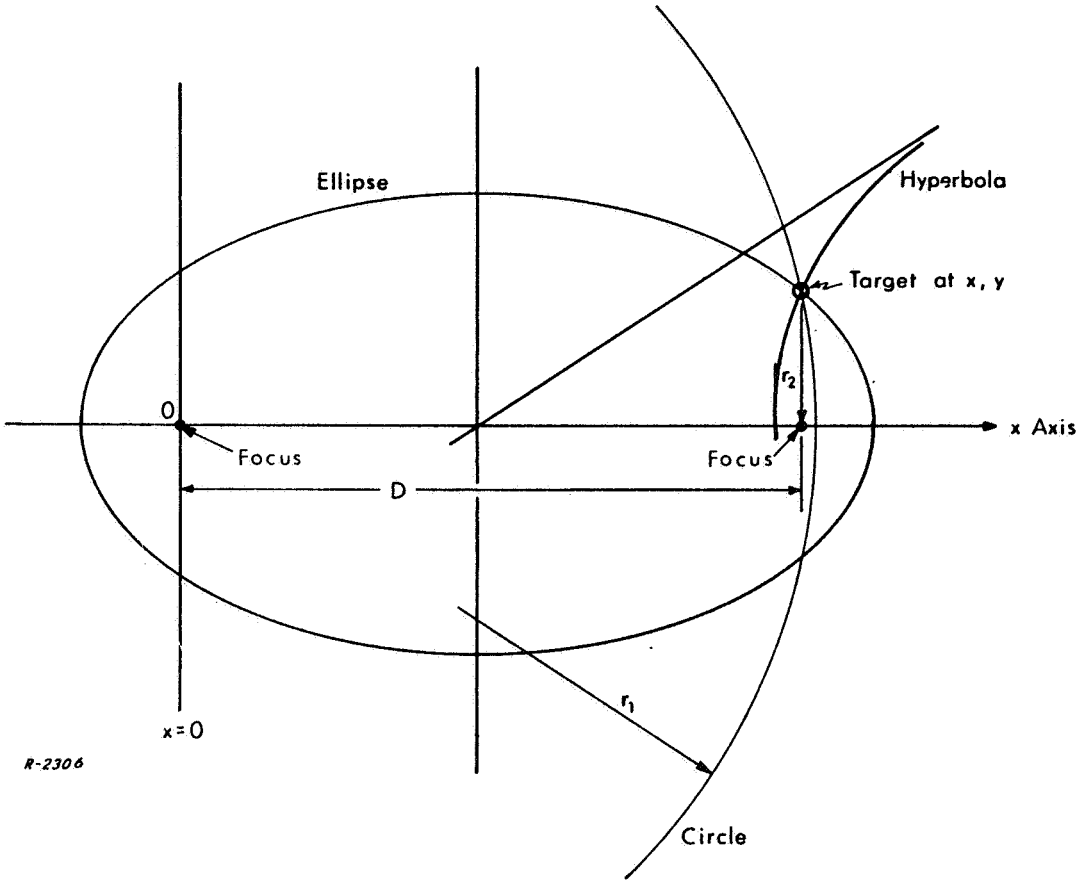


Fig. 19 Intersection of Circle, Hyperbola, Ellipse

with solution

$$x = (D^2 - R_d^2 \pm 2r_1R_d)/2D \tag{28a}$$

$$y = \left[r_1^2 - (D^2 - R_d^2 \pm 2r_1R_d)^2/4D^2\right]^{1/2} \tag{28b}$$

The solutions for the target position have the same form in both examples. Equation (26a) differs from Eq. (28a) only in that R_s is replaced by R_d, and similarly for Eq. (26b) and Eq. (28b). In both sets of equations the ± sign is resolved by taking the plus sign for an assumed target position above and to the right of the center of the ellipse and hyperbola. In actual practice the sign can be resolved by checking the relative lengths of r_1 and r_2.

The error in x and y given by Δx and Δy, respectively, may be found by assuming errors in r_1, R_s and R_d given by Δr_1, ΔR_s and ΔR_d respectively. It is clear from the symmetry of Eqs. (26) and (28) that the error for the case of the ellipse must have the same form as that for the case of the hyperbola upon substituting ΔR_d for ΔR_s. To determine the error it therefore suffices to calculate the error for the case of the ellipse only, since the hyperbola will give the same error.

Δx is obtained by replacing x with $x+\Delta x$, R_s with $R_s+\Delta R_s$, and r_1 with $r_1+\Delta r_1$ in Eq. (26a) giving

$$x + \Delta x = [D^2 - (R_s + \Delta R_s)^2 + 2(r_1 + \Delta r_1)(R_s + \Delta R_s)](1/2D) \qquad (29)$$

After expanding Eq. (29), Eq. (26a) is subtracted out giving

$$\Delta x = \frac{r_1 - R_s}{D} \Delta R_s + \frac{R_s}{D} \Delta r_1 \qquad (30)$$

Δy is obtained by a similar procedure. First Eq. (26b) is squared to remove the radical sign. Then y is replaced with $y+\Delta y$, R_s with $R_s+\Delta R_s$ and r_1 with Δr_1. The resulting equation is expanded and all higher order terms ($\overline{\Delta y}^2$ and $\overline{\Delta r_1}^i \overline{\Delta R_s}^j$ where $i + j \geq 2$) are dropped, leaving an expression for $y^2 + 2y\Delta y$. The squared Eq. (26b) is then subtracted out leaving an expression for $2y\Delta y$. Finally this expression is divided by Eq. (26b) to give

$$\Delta y = \frac{\left[\begin{array}{l}(4D^2R_s - 4D^2r_1 - 4R_s^3 + 12r_1R_s^2 - 8r_1^2R_s)\Delta R_s \\ \\ \quad + (4R_s^3 - 4D^2R_s - 8R_s^2r_1)\Delta r_1\end{array}\right]}{8D^2\sqrt{r_1^2 - (D^2 - R_s^2 + 2r_1R_s)^2(1/4D^2)}} \qquad (31)$$

These two examples show that the same error is obtained for the case of circle plus ellipse, or circle plus hyperbola, target position solutions and demonstrate the assertion that the GDOP depends only on the placing of the ground sites rather than on the surfaces which may be generated as a convenience in solving the problem.

Other possible combinations of surfaces which can be used for obtaining target positions are a set of ellipsoids, a set of both ellipsoids and hyperboloids, and a set of hyperboloids only. The latter presents an opportunity for passive position location as is discussed below. When a large number of measurements (i.e., ground stations) are involved, the question arises as to

the optimum method of combining data from all sources. In the discussion which follows, the least-squares data reduction technique is derived and presented with examples to demonstrate its application in position location systems.

The Least-Squares Data Reduction Technique

In this section, the least-squares data reduction technique is derived and presented. First, the case where many similar measurements are performed to determine just a few variables is considered. Then, the method is generalized to include dissimilar measurements.

Least-Squares Estimates of Parameters Based On a Series of Similar Measurements

The problem of estimating vehicle position from a series of range difference measurements is used in the following discussion, for illustrative purposes, to introduce least-squares methods.

Consider the two-dimensional position location problem shown in Fig. 20. Transmitting stations are located at (x_1, y_1), (x_2, y_2) ... (x_5, y_5) and the vehicle to be tracked is located at (x_t, y_t). Assume that four independent range differences can be measured at the vehicle, i.e., $\Delta r_1 = r_2 - r_1$, $\Delta r_2 = r_3 - r_2$, $\Delta r_3 = r_4 - r_3$, and $\Delta r_4 = r_3 - r_4$. Note that if only three transmitting sites were used giving only two independent range-differences, then a straightforward computation could be performed to determine x_t and y_t. The computation is similar to that described in the section entitled, "Utilization of Surfaces in Trajectory Situations," p. 37. When there are two equations and two unknowns, the estimates of target position are related uniquely to the measurements. On the other hand, when redundant measurements are available, as is the case depicted in Fig. 20, more than one possible solution for vehicle position exists. The problem, then, is to

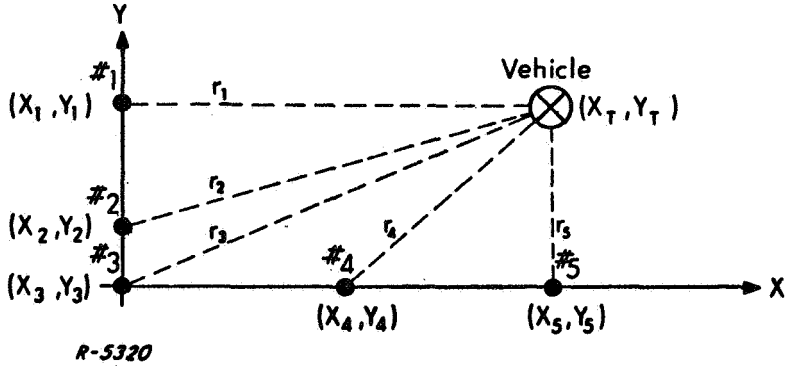


Fig. 20 Simple Position Determination Problem

optimally combine the redundant measured data in a way which accounts for individual measurement accuracy and the sensitivity of the measurements to errors as a function of target and station position. As noted above, the discussion of this section is restricted to the case where all measurements have errors described by the same statistics.

The equations which relate the measurements, Δr_i, to the vehicle position are

$$\begin{aligned}
\Delta r_1 &= r_2 - r_1 = [(x_2 - x_t)^2 + (y_2 - y_t)^2]^{1/2} - [(x_1 - x_t)^2 + (y_1 - y_t)^2]^{1/2} + \epsilon_1 \\
\Delta r_2 &= r_3 - r_2 = [(x_3 - x_t)^2 + (y_3 - y_t)^2]^{1/2} - [(x_2 - x_t)^2 + (y_2 - y_t)^2]^{1/2} + \epsilon_2 \\
\Delta r_3 &= r_4 - r_3 = [(x_4 - x_t)^2 + (y_4 - y_t)^2]^{1/2} - [(x_3 - x_t)^2 + (y_3 - y_t)^2]^{1/2} + \epsilon_3 \\
\Delta r_4 &= r_5 - r_4 = [(x_5 - x_t)^2 + (y_5 - y_t)^2]^{1/2} - [(x_4 - x_t)^2 + (y_4 - y_t)^2]^{1/2} + \epsilon_4
\end{aligned} \tag{32}$$

where ϵ_i is the error associated with the measurement of Δr_i. It is assumed that all the ϵ_i are mean zero random variables with variance, σ^2.

Equations of the form of Eq. (32) are handled with some difficulties when there are only two. However, when the simultaneous treatment of 10 or 20 such equations is considered, it is found best to use matrix notation and matrix algebra. To do this, it is necessary to linearize the equations. Linearization is accomplished with the aid of a Taylor series expansion of each of the equations about a preliminary estimate of the vehicle position (x_{to}, y_{to}). Assuming that the preliminary estimate of vehicle position is accurate enough, the series is truncated at the first term. Thus the i^{th} linearized equation is

$$\begin{aligned}
\Delta r_i(x_t, y_t) \approx \Delta r_i(x_{to}, y_{to}) &+ \left.\frac{\partial \Delta r_i(x_t, y_t)}{\partial x_t}\right|_{(x_{to}, y_{to})} (x_t - x_{to}) \\
&+ \left.\frac{\partial \Delta r_i(x_y, y_t)}{\partial y_t}\right|_{(x_{to}, y_{to})} (y_t - y_{t_o}) + \epsilon_i
\end{aligned} \tag{33}$$

This, in turn, can be simplified by the following changes in notation: define $\mu_i = \Delta r_i(x_t, y_t) - \Delta r_i(x_{to} - y_{to})$, i.e., the difference between the actual measurement value and what would have been measured in the absence of noise if the vehicle was actually located at (x_{to}, y_{to}). Similarly, define position corrections

$$\beta_1 = (x_t - x_{to})$$

$$\beta_2 = (y_t - y_{to})$$

and coefficients

$$a_{1i} = \left.\frac{\partial \Delta r_i(x_t, y_t)}{\partial x_t}\right|_{(x_{to}, y_{to})}$$

$$a_{2i} = \left.\frac{\partial \Delta r_i(y_t, y_t)}{\partial y_t}\right|_{(x_{to}, y_{to})} \tag{34}$$

Thus, Eq. (32) can be written as

$$\begin{aligned}
\mu_1 &= a_{11}\beta_1 + a_{21}\beta_2 + \epsilon_1 \\
\mu_2 &= a_{12}\beta_1 + a_{22}\beta_2 + \epsilon_2 \\
\mu_3 &= a_{13}\beta_1 + a_{23}\beta_2 + \epsilon_3 \\
\mu_4 &= a_{14}\beta_1 + a_{24}\beta_2 + \epsilon_4
\end{aligned} \tag{35}$$

Finally, using matrix notation, Eq. (35) becomes

$$[\mu] = [a_t][\beta] + [\epsilon] \tag{36}$$

The least-squares estimates of the β_i are found in a straightforward manner. Specifically, the sum of errors squared is minimized by differentiating the sum squared error with respect to the β_i and setting the derivatives equal to zero.

$$\frac{\partial\left(\sum_{i}^{m} \epsilon_i^2\right)}{\partial \beta_k} = \frac{\partial}{\partial \beta_k} \sum_{i}^{m} \left(\mu_i - \sum_{j}^{n} a_{ji}\beta_j\right)^2 = 0$$

$$= -2 \sum_{i}^{m} \left(\mu_i - \sum_{j}^{n} a_{ji}\beta_j\right) a_{ki} = 0 \qquad (37)$$

$$k=1, 2, \ldots, n$$

where m is the number of equations and n is the number of unknowns.

Recasting Eq. (37) in matrix notation yields

$$a\mu - aa_t\beta = 0 \qquad (38)$$

Thus, the least-squares parameter estimates are given by

$$\beta = [a \cdot a_t]^{-1} a\mu = \alpha\mu \qquad (39)$$

where α has been implicitly defined. Notice that for the special case when a is a square matrix, α reduces to a_t^{-1}. This can be seen by expanding Eq. (39) as follows:

$$\beta = a_t^{-1} a^{-1} a\mu = a_t^{-1} \mu \qquad (40)$$

This is the solution of Eq. (36) obtainable in a straightforward manner for the case of n equations and n unknowns.

Next, it is desired to determine the variances in estimation of the β_i because these can be directly related to position location errors.

To this end, it is noted that the covariance matrix of μ, i.e.,

$$\text{cov}(\mu) = M = \begin{bmatrix} \sigma_{\mu_1}^2 & 0 & 0 & 0 \\ 0 & \sigma_{\mu_2}^2 & \vdots & \vdots \\ \vdots & \vdots & \vdots & \sigma_{\mu_n}^2 \end{bmatrix} \qquad (41)$$

is given by

$$M = \overline{(\mu - \bar{\mu})(\mu - \bar{\mu})_t} \tag{42}$$

where the "bar" denotes average value. But, from Eq. (36), $\bar{\mu} = a_t\beta$. Thus, Eq. (42) becomes

$$\overline{(\mu - \bar{\mu})(\mu - \bar{\mu})_t} = \overline{\epsilon\epsilon}_t \tag{43}$$

It is recalled that for the problem under consideration here, the errors ϵ_i are all mean zero random variables with the same variance σ^2. Therefore, Eq. (43) becomes

$$\overline{(\mu - \bar{\mu})(\mu - \bar{\mu})_t} = \sigma^2 I \tag{44}$$

where I is the identity matrix. Now, the covariance matrix of the position corrections β_i may be expressed in terms of the μ_i with the aid of Eq. (39) as

$$\text{cov}(\beta) = B = \overline{(\beta - \bar{\beta})(\beta - \bar{\beta})_t} = \alpha\overline{(\mu - \bar{\mu})(\mu - \bar{\mu})_t}\,\alpha_t \tag{45}$$

or, substituting from Eq. (44)

$$B = \sigma^2\alpha \cdot \alpha_t = \sigma^2(a \cdot a_t)^{-1}\, a[(a \cdot a_t)^{-1}\, a]_t$$

$$= \sigma^2(a \cdot a_t)^{-1}\, a \cdot a_t(a \cdot a_t)_t^{-1} = \sigma^2(a \cdot a_t)^{-1} \tag{46}$$

where use has been made of the combinational properties of matrices and the fact that $a \cdot a_t$ is symmetrical.

The variance of the position corrections B_{11} and B_{22} are the major diagonal elements of cov(β). It is further recognized that the variances in the position corrections are equal to the variances in the coordinates x and y. Thus, the total position error for the problem defined by Fig. 20 and Eq. (32) is given by

$$\sigma_p^2 = B_{11} + B_{22} \tag{47}$$

The derivation just presented is easily generalized to three dimensions. For this case, Eq. (32) will have the form

$$\Delta r_i = r_j - r_i = \left[(x_j - x_t)^2 + (y_j - y_t)^2 + (z_j - z_t)^2\right]^{1/2}$$

$$-\left[(x_i - x_t)^2 + (y_i - y_t)^2 + (z_i z_t)^2\right]^{1/2} + \epsilon_i \qquad (48)$$

Of course, there must now be a minimum of three such equations to solve for three unknowns x_t, y_t and z_t. After linearization, Eq. (48) will become

$$\mu_i = a_{1i}\beta_1 + a_{2i}\beta_2 + a_{3i}\beta_3 + \epsilon_i$$

which corresponds in form to Eq. (35). The manipulations required for the three dimensional solution are identical to those described by Eq. (36) through Eq. (46). Finally, the position error will be

$$\sigma_p^2 = B_{11} + B_{22} + B_{33}$$

which is the three dimensional version of Eq. (46). The additional term B_{33} is of course the variance in z.

Least-Squares Combining of Data From a Variety of Measurements

The special case considered above assumed that the measurement errors, ϵ_i, were independent mean zero random variables with variance σ^2. This section discusses a more general problem in which the ϵ_i are again independent mean zero random variables but the variances of all the ϵ_i are not necessarily equal.

Such a situation does arise when several dissimilar measurements are employed to determine position. For example, the set of range difference measurements of Eq. (32) may be augmented by a range-angle measurement to an extra station. The geometry of this situation is shown in Fig. 21. The equation for r_6 is similar in form to Eq. (32), but θ is substantially different. Specifically,

$$r_6 = \sqrt{(x_6 - x_t)^2 + (y_6 - y_t)^2} \qquad (49)$$

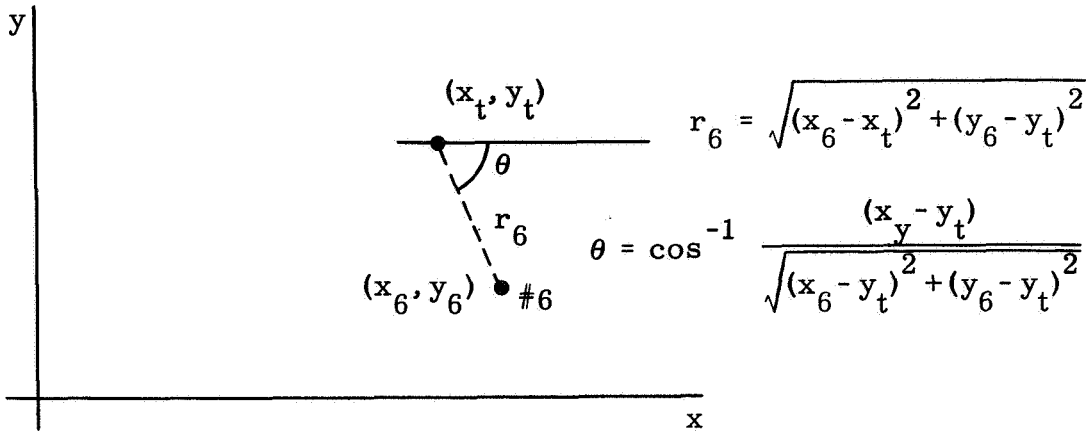


Fig. 21 Arbitrary Range-Angle Measurement (in 2 Dimensions)

and

$$\theta = \cos^{-1} \frac{(x_6 - x_t)}{\sqrt{(x_6 - x_t)^2 + (y_6 - y_t)^2}} \tag{50}$$

The errors associated with these two measurements are reasonably assumed to have different statistical descriptions because the measurements performed are completely different.

The fact that the equations describing the range-angle measurements are different in form than those of Eq. (32) is of little concern because they will be linearized anyway. In this way, six equations of the form of Eq. (35) are generated. The first four equations are in fact given in Eq. (35). The fifth equation is the linearized version of Eq. (49) and the sixth is the linearized version of Eq. (50). The problem, then, is to find the least-squares estimates of β_1 and β_2 where

$$\begin{aligned}
\mu_1 &= a_{11}\beta_1 + a_{21}\beta_2 + \epsilon_1 \\
\mu_2 &= a_{12}\beta_1 + a_{22}\beta_2 + \epsilon_2 \\
\mu_3 &= a_{13}\beta_1 + a_{23}\beta_2 + \epsilon_3 \\
\mu_4 &= a_{14}\beta_1 + a_{24}\beta_2 + \epsilon_4 \\
\mu_5 &= a_{15}\beta_1 + a_{25}\beta_2 + \epsilon_5 \\
\mu_6 &= a_{16}\beta_1 + a_{26}\beta_2 + \epsilon_6
\end{aligned} \tag{51}$$

In summary, the only difference between the problem considered here and the one discussed in the section entitled, "Least Squares Estimates of Parameters Based on a Series of Similar Measurements," p. 45, is that Eq. (44), i.e., $\overline{\epsilon\epsilon}_t = \sigma^2 I$, does not hold. Instead, the error covariance matrix is given by

$$\overline{\epsilon\epsilon}_t = \begin{bmatrix} \sigma_\Delta^2 & & & & & & \\ & \sigma_\Delta^2 & & & & \bigcirc & \\ & & \sigma_\Delta^2 & & & & \\ & & & \sigma_\Delta^2 & & & \\ & & & & \sigma_r^2 & & \\ & \bigcirc & & & & & \\ & & & & & & \sigma_\theta^2 \end{bmatrix} \tag{52}$$

where $\overline{\epsilon\epsilon}_t = \overline{(\mu - \mu)(\mu - \mu)_t} = \text{cov}(\mu) = M$, as before, σ_Δ^2 is the variance of the range difference measurements, σ_r^2 is the variance of the straight range measurement and σ_θ^2 is the variance of the angle measurement. For this case, the least squares estimates of the β_i are found by minimizing the weighted sum of errors squared.* Thus, Eq. (37) is replaced by

$$\frac{\partial \sum_i \omega_i \epsilon_i^2}{\partial \beta_k} = \frac{\partial \sum_i}{\partial \beta_k} \omega_i \left(\mu_i - \sum_j a_{ji} \beta_j \right)^2 = 0 \tag{53}$$

where the optimum weighting coefficients, ω_i, are given by

$$[\omega] = K[M]^{-1} \tag{54}$$

and K is a scalar multiplier chosen so that

$$\sum_i K\omega_i = 1 \tag{55}$$

* Scheffe, H., "The Analysis of Variance," Wiley, 1959, p. 20.

The weighting factors for the example considered here are

$$\omega_1 = \omega_2 = \omega_3 = \omega_4 = K/\sigma_\Delta^2$$
$$\omega_5 = K/\sigma_r^2 \tag{56}$$
$$\omega_6 = K/\sigma_\theta^2$$

and

$$K = \frac{1}{\frac{4}{\sigma_\Delta^2} + \frac{1}{\sigma_r^2} + \frac{1}{\sigma_\theta^2}} \tag{57}$$

Carrying out the indicated operations of Eq. (53) yields a set of equations

$$-2 \sum_i^m \omega_i \left(\mu_i - \sum_j^n a_{ji} \beta_j \right) a_{ki} \qquad k=1, 2, \cdots, \mu$$

or

$$\sum_i^m \omega_i \mu_i a_{ki} = \sum_i^m \sum_j^n \omega_i a_{ji} a_{ki} \beta_j \qquad k=1, 2, \cdots, n \tag{58}$$

Equation (58) may be expressed in matrix notation as

$$a\omega\mu = (a\omega a_t)\beta \tag{59}$$

Finally, the least-squares estimates of the β_i are found to be

$$\beta = (a\omega a_t)^{-1} a\omega\mu = \gamma\mu \tag{60}$$

Now, the variances in the estimates may be computed following the procedure in the section entitled, "Least Squares Estimates of Parameters Based on a Series of Similar Measurements," p. 45. The covariance matrix of the position corrections β_i is given by

$$\text{cov}(\beta) = B = \overline{(\beta - \bar{\beta})(\beta - \bar{\beta})_t} = \gamma \overline{(\mu - \bar{\mu})(\mu - \bar{\mu})_t} \gamma_t \tag{61}$$

where γ is defined in Eq. (60) and $\overline{(\mu - \Pi)(\mu - \Pi)}$ is defined in Eq. (52). Substituting from Eqs. (52) and (60) yields,

$$B = (a\omega a_t)^{-1} a\omega\overline{(e \cdot e_t)}\left((a\omega a_t)^{-1} a\omega\right)_t \tag{62}$$

Equation (62) is simplified by noting the definitions of ω and $\overline{e \cdot e_t}$. Namely, from Eq. (54),

$$\omega\overline{(e \cdot e_t)} = K[M]^{-1} M = K \tag{63}$$

a scalar. Then, noting that ω is a diagonal matrix,

$$B = K(a\omega a_t)^{-1} a\omega_t a_t\left((a\omega a_t)^{-1}\right)_t$$

which reduces to

$$= K(a\omega a_t)^{-1} \tag{64}$$

since $(a\omega a_t)^{-1}$ is symmetrical.

As before, the major diagonal elements of B are the variances of the position corrections, β. That is B_{11} is the variance of β_1, B_{22} is the variance of β_2, etc.

Investigation of GDOP Phenomena Using the Least-Squares Data Reduction Technique

The least-squares data reduction technique has been employed to investigate GDOP for two distinct configurations. First, the GDOP errors as a function of vehicular and station position are computed for a hyperbolic navigation system. The problem is solved for the two dimensional case, e.g., lat. long. A modified form of the hyperbolic navigation system is considered in which use is made of a one-way ranging capability on-board the vehicle. The second GDOP problem considers the errors in altitude of a trilateration or multilateration altimeter. The altimeter may be realized using DME/DME/DME or via a new configuration employing one-way ranging and time reference systems.

GDOP in Hyperbolic and Augmented Hyperbolic Navigation Systems

A computer program was written based on Eq. (64) to determine the error as a function of vehicle and station location for a typical hyperbolic

navigation system and for a hyperbolic navigation system augmented by a one-way range measurement obtained with the aid of an on-board stable time reference system.

The computer investigation consisted of 42 computer runs. Each of these runs considered a slightly different configuration. The results of each run consist of the rms error in position location for 36 different vehicle positions defined by x = 200, 600, 1000, 1400, 1800, 2200 and y = 200, 600, 1000, 1400, 1800, 2200. Six different station configurations were considered. These are listed in Table I. The dimension of the vehicle and station positions may be arbitrarily assigned as long as the particular units selected are consistently used for both vehicle and station positions.

Hyperbolic and augmented hyperbolic navigation is implemented for each of the six station configurations. Three measurements are performed while only two are sufficient for each of the techniques. For the hyperbolic case, the three measurements are range differences among the four stations. The variances of the first two difference measurements are normalized to unity while the redundant third measurement has variance of 1, 10, or 100. Similarly, for the augmented hyperbolic case, two range differences and a one-way range measurement (to station #2) are implemented. Station #4 is not used in this case. Again, the variances of the first two range differences are normalized to unity while the redundant one-way range measurement takes on one of four values for its variance, i.e., 1, 10, 100 and 1000. Thus, there are forty-two distinct configurations and computer runs. These are shown in Table II.

TABLE I

STATION CONFIGURATIONS

	I	II	III	IV	V	VI
#1	x = 0 y = 400	x = 0 y = 800	x = 0 y = 1600	x = 0 y = 400	x = 0 y = 800	x = 0 y = 1600
#2	x = 0 x = 0	x = 0 x = 0	x = 0 x = 0	x = 0 x = 0	x = 0 x = 0	x = 0 x = 0
#3	x = 200 y = 0	x = 400 y = 0	x = 800 y = 0	x = 0 y = -400	x = 0 y = -800	x = 0 y = -1600
#4	x = 400 y = 0	x = 800 y = 0	x = 1600 y = 0	x = 400 y = 0	x = 800 y = 0	x = 1600 y = 0

TABLE II

SYSTEM CONFIGURATIONS

Run Number	Station Configuration	Navigation Technique	Redundant Measurement Variance
1-6	I-VI	hyperbolic	1
7-12	I-VI	hyperbolic	10
13-18	I-VI	hyperbolic	100
19-24	I-VI	hyperbolic	1
25-30	I-VI	aug. hyp.	10
31-36	I-VI	aug. hyp.	100
37-42	I-VI	aug. hyp.	1000

The results of these computations are discussed below. Each run produces a grid of numbers. These are the rms position errors as a function of vehicle position, i.e.,

$$\text{rms position error} = [\sigma_x^2 + \sigma_y^2]^{1/2} \tag{65}$$

The orientation of the coordinate system in the computer printouts is shown graphically in Fig. 22. Figure 22 also shows, for the purposes of example, the location of stations in configuration III.

Results of the Hyperbolic Navigation Investigation

In this section, the results of several runs are presented side by side for ease of comparison. The significance of each comparison and salient features of each of the cited runs are described below each pair of runs.

Figure 23 is a comparison of run #1 vs run #3. It shows the advantage of a large station array vs a relatively small station array, all other conditions being equal. Notice that the position error increases as the vehicle position moves away from the cluster of stations and along the axis on which the stations are located. This is as expected from the analysis in the section entitled, "Utilization of Surfaces in Trajectory Situations," p. 37. The

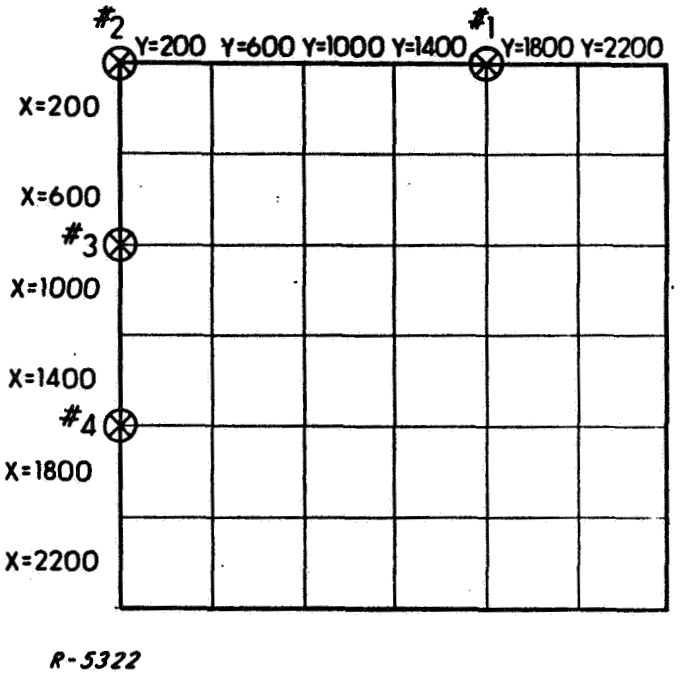


Fig. 22 Orientation of Computer Printout Coordinate System with Station Configuration III Superimposed

9.42E-01	5.09E+00	3.01E+01	8.11E+01	1.53E+02	2.44E+02
5.06E+00	6.82E+00	1.73E+01	4.16E+01	8.49E+01	1.50E+02
3.15E+01	1.70E+01	2.36E+01	4.03E+01	6.96E+01	1.15E+02
1.02E+02	4.04E+01	3.95E+01	5.16E+01	7.44E+01	1.09E+02
2.38E+02	8.37E+01	6.73E+01	7.30E+01	9.07E+01	1.19E+02
4.61E+02	1.54E+02	1.11E+02	1.06E+02	1.18E+02	1.41E+02

Fig. 23a Run #1

8.35E-01	8.51E-01	9.57E-01	1.11E+00	2.98E+00	1.01E+01
8.62E-01	8.28E-01	9.49E-01	1.24E+00	2.10E+00	4.25E+00
1.01E+00	9.85E-01	1.15E+00	1.51E+00	2.24E+00	3.66E+00
1.24E+00	1.31E+00	1.53E+00	1.95E+00	2.67E+00	3.86E+00
3.04E+00	2.16E+00	2.28E+00	2.68E+00	3.38E+00	4.49E+00
1.06E+01	4.26E+00	3.67E+00	3.87E+00	4.48E+00	5.50E+00

Fig. 23b Run #3

position error magnitudes are normalized to the measurement error units. In particular, the errors in measurement for both run #1 and run #3 are set equal to 1. If this is assumed to represent variances of 1 ft^2, then the position errors are given in "feet." If the actual measurement error is, say, 10 ft (i.e., variance of 100) then each position error printed out should be multiplied by 10.

Examination of run #1 vs run #3 shows that a large station array greatly reduces the GDOP. This improvement occurs because the Lines-of-Position, in the large array case, do not intersect at very shallow angles; while, in the small array case, the Lines-of-Position will intersect at shallow angles when the vehicle is far from the station array.

A comparison of run #1 vs run #7, Fig. 24, shows the relative improvement obtained by the redundant range difference measurement. In run #7, the variance of the redundant measurement is increased from 1 to 10 (i.e., rms error of the redundant measurement is 3.16 units). This has the effect of roughly doubling the rms error in regions of poor GDOP while leaving the rms error in good GDOP regions unchanged.

This effect is more clearly demonstrated in Fig. 25 in which runs #3 and #9 are compared. The comparison is the same as that of run #1 vs #7 except that station configuration III is assumed instead of configuration I. Note, again, that the latter configuration produces a very large region of poor GDOP.

The effect of a poor quality redundant measurement is further demonstrated in Fig. 26 in which run #3 is compared to run #15. Again, station configuration III is assumed, but the rms error of the redundant measurement is assumed to be 10 units. Notice that many of the position errors in the good GDOP region are at most doubled. This is the result of a property of least-squares combining in which measurements of known inferior quality are given less weight in the position determination process.

Figure 27 shows a comparison of run #1 with run #19. Run #19 is the same as run #1 except that the redundant measurement is obtained by one-way ranging to station #2. The variances of all measurements in runs #1 and #19 are set to unity. Note the dramatic improvement in accuracy at distances far from the station array. In fact, the point at which the results of Fig. 27 are inferior is at x = 200, y = 200. Elsewhere the GDOP is reduced by more than an order of magnitude.

Figure 28 shows a comparison similar to that of Fig. 27 except that station configuration III is considered instead of configuration I. In this comparison, the improvement in regions of poor GDOP is still large but less than an order of magnitude. At the same time, the augmented

9.42E-01	5.09E+00	3.01E+01	8.11E+01	1.53E+02	2.44E+02	1.40E+00	7.94E+00	4.35E+01	1.22E+02	2.38E+02	3.89E+02
5.06E+00	6.82E+00	1.73E+01	4.16E+01	8.49E+01	1.50E+02	1.12E+01	1.24E+01	2.81E+01	6.28E+01	1.24E+02	2.18E+02
3.15E+01	1.70E+01	2.36E+01	4.03E+01	6.96E+01	1.15E+02	6.97E+01	3.34E+01	4.20E+01	6.64E+01	1.08E+02	1.72E+02
1.02E+02	4.04E+01	3.95E+01	5.16E+01	7.44E+01	1.09E+02	2.17E+02	8.19E+01	7.43E+01	9.06E+01	1.23E+02	1.73E+02
2.38E+02	8.37E+01	6.73E+01	7.30E+01	9.07E+01	1.19E+02	4.94E+02	1.71E+02	1.31E+02	1.34E+02	1.58E+02	1.99E+02
4.61E+02	1.54E+02	1.11E+02	1.06E+02	1.18E+02	1.41E+02	9.40E+02	3.14E+02	2.14E+02	2.00E+02	2.12E+02	2.44E+02

Fig. 24a Run #1

Fig. 24b Run #7

8.35E-01	8.51E-01	9.57E-01	1.11E+00	2.98E+00	1.01E+01	9.33E-01	1.08E+00	1.37E+00	1.76E+00	4.50E+00	1.42E+01
8.62E-01	8.28E-01	9.49E-01	1.24E+00	2.10E+00	4.25E+00	9.42E-01	1.07E+00	1.42E+00	2.01E+00	3.43E+00	6.57E+00
1.01E+00	9.85E-01	1.15E+00	1.51E+00	2.24E+00	3.66E+00	1.62E+00	1.55E+00	1.89E+00	2.55E+00	3.79E+00	6.00E+00
1.24E+00	1.31E+00	1.53E+00	1.95E+00	2.67E+00	3.86E+00	2.06E+00	2.40E+00	2.79E+00	3.49E+00	4.68E+00	6.59E+00
3.04E+00	2.16E+00	2.28E+00	2.68E+00	3.38E+00	4.49E+00	7.29E+00	4.47E+00	4.46E+00	5.04E+00	6.16E+00	7.92E+00
1.06E+01	4.26E+00	3.67E+00	3.87E+00	4.48E+00	5.50E+00	2.89E+01	9.68E+00	7.63E+00	7.60E+00	8.45E+00	1.00E+01

Fig. 25a Run #3

Fig. 25b Run #9

8.35E-01	8.51E-01	9.57E-01	1.11E+00	2.98E+00	1.01E+01
8.62E-01	8.28E-01	9.49E-01	1.24E+00	2.10E+00	4.25E+00
1.01E+00	9.85E-01	1.15E+00	1.51E+00	2.24E+00	3.66E+00
1.24E+00	1.31E+00	1.53E+00	1.95E+00	2.67E+00	3.86E+00
3.04E+00	2.16E+00	2.28E+00	2.68E+00	3.38E+00	4.49E+00
1.06E+01	4.26E+00	3.67E+00	3.87E+00	4.48E+00	5.50E+00

Fig. 26a Run #3

9.51E-01	1.13E+00	1.48E+00	1.94E+00	5.18E+00	1.73E+01
9.59E-01	1.14E+00	1.57E+00	2.30E+00	4.02E+00	7.86E+00
2.36E+00	1.89E+00	2.27E+00	3.08E+00	4.61E+00	7.34E+00
5.08E+00	3.89E+00	3.82E+00	4.52E+00	5.95E+00	8.30E+00
1.78E+01	8.06E+00	6.68E+00	6.93E+00	8.13E+00	1.02E+01
4.93E+01	1.62E+01	1.15E+01	1.07E+01	1.14E+01	1.32E+01

Fig. 26b Run #15

9.42E-01	5.09E+00	3.01E+01	8.11E+01	1.53E+02	2.44E+02	1.20E+00	2.08E+00	4.42E+00	6.66E+00	8.77E+00	1.08E+01
5.06E+00	6.82E+00	1.73E+01	4.16E+01	8.49E+01	1.53E+02	1.93E+00	2.29E+00	3.59E+00	5.53E+00	7.63E+00	9.80E+00
3.15E+01	1.70E+01	2.36E+01	4.03E+01	6.96E+01	1.15E+02	2.77E+00	3.02E+00	3.88E+00	5.28E+00	7.03E+00	9.00E+00
1.02E+02	4.04E+01	3.95E+01	5.16E+01	7.44E+01	1.09E+02	3.70E+00	3.88E+00	4.53E+00	5.60E+00	7.32E+00	8.70E+00
2.38E+02	8.37E+01	6.73E+01	7.30E+01	9.07E+01	1.19E+02	4.66E+00	4.81E+00	5.32E+00	6.18E+00	7.36E+00	8.78E+00
4.61E+02	1.54E+02	1.11E+02	1.06E+02	1.18E+02	1.41E+02	5.63E+00	5.75E+00	6.18E+00	6.93E+00	7.89E+00	9.12E+00

Fig. 27a Run #1

Fig. 27b Run #19

8.35E-01	8.51E-01	9.57E-01	1.11E+00	2.98E+00	1.01E+01	9.33E-01	1.10E+00	1.42E+00	1.58E+00	1.66E+00	2.60E+00
8.62E-01	8.28E-01	9.49E-01	1.24E+00	2.10E+00	4.25E+00	9.10E-01	1.02E+00	1.25E+00	1.38E+00	1.51E+00	1.93E+00
1.01E+00	9.85E-01	1.15E+00	1.51E+00	2.24E+00	3.66E+00	1.35E+00	1.26E+00	1.34E+00	1.43E+00	1.56E+00	1.83E+00
1.24E+00	1.31E+00	1.53E+00	1.95E+00	2.67E+00	3.36E+00	1.64E+00	1.51E+00	1.49E+00	1.54E+00	1.66E+00	1.87E+00
3.94E+00	2.16E+00	2.28E+00	2.68E+00	3.38E+00	4.49E+00	1.78E+00	1.68E+00	1.65E+00	1.69E+00	1.79E+00	1.97E+00
7.06E+01	4.26E+00	3.67E+00	3.87E+00	4.48E+00	5.50E+00	1.94E+00	1.86E+00	1.83E+00	1.86E+00	1.95E+00	2.11E+00

Fig. 28a Run #3

Fig. 28b Run #21

hyperbolic scheme produces results slightly inferior to those of the straight hyperbolic method in regions of good GDOP.

Figure 29 shows run #21 along with runs #27, #33 and #39. These latter are the same as run #21 except that the variance of the redundant one-way range measurement is increased to 10, 100, and 1000, respectively. Comparing these results with those of run #3 shows that as the variance of the one-way redundant measurement is increased, there is a point at which it is better, in general, to use the straight hyperbolic navigation technique.

Figure 30 is the same as Fig. 29 except that station configuration I is assumed. It may be concluded, by examination of Figs. 29 and 30, that the "cross-over" point at which the straight hyperbolic navigation is superior to augmented hyperbolic navigation is a strong function of the GDOP characteristics of the system. That is, in Fig. 30, the augmented system continues to provide improved performance for the majority of vehicle positions even when the variance of the redundant measurement is assumed to be 1000 times that of the range difference measurements. On the other hand, Fig. 29 shows that when GDOP is not severe, the augmented system has deteriorated performance when the redundant measurement has large variance.

GDOP in a Multilateration Altimeter

A computer program was written to determine the error in latitude, longitude, and altitude of a multilateration navigation system. The rms errors in latitude, longitude, and altitude were printed out individually as a function of vehicle and station position.

9.33E-01	1.10E+00	1.42E+00	1.58E+00	1.66E+00	2.60E+00
9.10E-01	1.02E+00	1.25E+00	1.38E+00	1.51E+00	1.93E+00
1.35E+00	1.26E+00	1.34E+00	1.43E+00	1.56E+00	1.83E+00
1.64E+00	1.51E+00	1.49E+00	1.54E+00	1.66E+00	1.87E+00
1.78E+00	1.68E+00	1.65E+00	1.69E+00	1.79E+00	1.97E+00
1.94E+00	1.86E+00	1.83E+00	1.86E+00	1.95E+00	2.11E+00

Fig. 29a Run #21

9.51E-01	1.13E+00	1.48E+00	1.90E+00	3.03E+00	3.94E+00
9.55E-01	1.13E+00	1.53E+00	2.08E+00	2.79E+00	3.38E+00
2.17E+00	1.78E+00	2.05E+00	2.48E+00	2.94E+00	3.32E+00
3.54E+00	2.89E+00	2.79E+00	2.95E+00	3.17E+00	3.40E+00
3.67E+00	3.45E+00	3.30E+00	3.30E+00	3.39E+00	3.52E+00
3.71E+00	3.63E+00	3.54E+00	3.52E+00	3.55E+00	3.64E+00

Fig. 29b Run #27

9.53E-01	1.13E+00	1.49E+00	1.96E+00	4.76E+00	9.05E+00
9.60E-01	1.15E+00	1.59E+00	2.31E+00	3.87E+00	6.43E+00
2.51E+00	1.92E+00	2.30E+00	3.07E+00	4.38E+00	6.19E+00
7.80E+00	4.10E+00	3.82E+00	4.36E+00	5.39E+00	6.71E+00
1.01E+01	7.20E+00	6.07E+00	6.10E+00	6.68E+00	7.53E+00
1.05E+01	9.26E+00	8.17E+00	7.81E+00	7.97E+00	8.39E+00

Fig. 29c Run #33

9.54E-01	1.13E+00	1.49E+00	1.97E+00	5.22E+00	1.56E+01
9.61E-01	1.15E+00	1.59E+00	2.34E+00	4.09E+00	7.85E+00
2.56E+00	1.94E+00	2.33E+00	3.16E+00	4.71E+00	7.39E+00
1.03E+01	4.35E+00	4.02E+00	4.68E+00	6.08E+00	8.34E+00
2.16E+01	9.16E+00	7.10E+00	7.17E+00	8.26E+00	1.02E+01
2.88E+01	1.60E+01	1.18E+01	1.08E+01	1.13E+01	1.27E+01

Fig. 29d Run #39

1.20E+00	2.98E+00	4.42E+00	6.66E+00	8.77E+00	1.08E+01
1.93E+00	2.29E+00	3.59E+00	5.53E+00	7.63E+00	9.80E+00
2.77E+00	3.02E+00	3.88E+00	5.28E+00	7.03E+00	9.00E+00
3.70E+00	3.88E+00	4.53E+00	5.60E+00	7.02E+00	8.70E+00
4.66E+00	4.81E+00	5.32E+00	6.18E+00	7.36E+00	8.78E+00
5.63E+00	5.75E+00	6.18E+00	6.90E+00	7.89E+00	9.12E+00

Fig. 30a Run #19

1.53E+00	3.53E+00	5.35E+00	7.31E+00	9.27E+00	1.12E+01
3.64E+00	3.76E+00	4.67E+00	6.26E+00	8.20E+00	1.02E+01
4.13E+00	4.27E+00	4.91E+00	6.07E+00	7.65E+00	9.48E+00
4.78E+00	4.92E+00	5.44E+00	6.36E+00	7.64E+00	9.20E+00
5.55E+00	5.67E+00	6.11E+00	6.88E+00	7.94E+00	9.28E+00
6.38E+00	6.49E+00	6.88E+00	7.53E+00	8.45E+00	9.68E+00

Fig. 30b Run #25

1.60E+00	7.24E+00	1.08E+01	1.20E+01	1.33E+01	1.47E+01
9.30E+00	8.95E+00	1.03E+01	1.13E+01	1.25E+01	1.40E+01
1.05E+01	1.03E+01	1.06E+01	1.12E+01	1.22E+01	1.34E+01
1.07E+01	1.07E+01	1.09E+01	1.14E+01	1.22E+01	1.32E+01
1.10E+01	1.11E+01	1.13E+01	1.17E+01	1.24E+01	1.33E+01
1.15E+01	1.15E+01	1.17E+01	1.21E+01	1.27E+01	1.35E+01

Fig. 30c Run #31

1.61E+00	9.52E+00	2.83E+01	3.20E+01	3.28E+01	3.34E+01
1.64E+01	1.51E+01	2.43E+01	3.00E+01	3.20E+01	3.29E+01
3.07E+01	2.64E+01	2.79E+01	3.02E+01	3.16E+01	3.26E+01
3.19E+01	3.08E+01	3.05E+01	3.11E+01	3.18E+01	3.25E+01
3.21E+01	3.18E+01	3.16E+01	3.18E+01	3.21E+01	3.26E+01
3.22E+01	3.21E+01	3.21E+01	3.22E+01	3.24E+01	3.28E+01

Fig. 30d Run #37

As in the previous calculation both the measurement errors and the vehicle positions are given in normalized units. Figure 31 shows the format of the computer printouts. Each computer run places the vehicle at 54 positions x = 0, 40, ..., 200 and y = 0, 40, ..., 320. Three computer runs were performed for three hypothesized vehicle altitudes namely z = 2, 10, and 50 units. Four stations were assumed to be located as shown in Fig. 31 for a Multilateration Altimeter analysis. Four ranging stations are assumed at x = 40, y = 150, z = 0, at x = 40, y = 40, z = 0, at x = 180, y = 40, z = 0 and at x = 80, y = 80, and z = 0.2

The assumed navigation technique consists of measuring range from the vehicle to each of the four stations and then, based on the four range measurements, computing position in three dimensions. The variance of each of the range measurements is normalized to unity. Thus, as was the case above, the rms position errors printed out are similarly normalized. This calculation differs from that described in the section entitled, "Results of the Hyperbolic Navigation Investigation," p. 56, in that the position errors in the x direction, the y direction and the z direction are printed out separately instead of being combined root sum squared before printout. This was done so that the altitude error as a function of system geometry could be readily investigated. The three position errors are printed out, for all positions, as shown in x = 0, y = 0 position of Fig. 31.

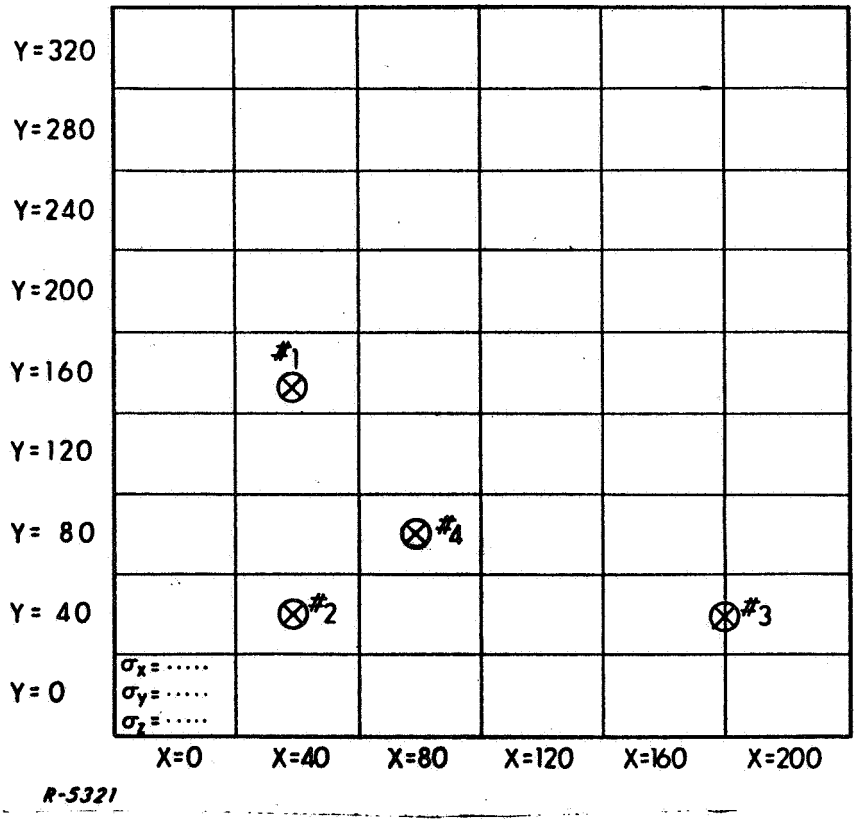


Fig. 31 Format of the Computer Printout

Results of Multilateration Altimeter Analysis

Results for the three altitudes z = 2, 10, and 50 are shown in Figs. 32a, 32b and 32c respectively.

It is convenient to interpret the units of x, y and z in two ways. First, they may be considered directly as corresponding to nautical miles. Then, the overall printout dimensions are 200 × 320 mm which is useful for consideration of, say, intercity navigation problems. In this case, only Figs. 32a and 32b are relevant since these correspond to altitudes of 2 nm = 12,000 ft and 10 nm = 60,000 ft, respectively. Alternatively, the printouts may be used to analyze a multilateration approach system by interpreting the units of x, y, and z as multipliers of 100 ft. That is, x = 40 corresponds to x = 4000 ft, etc. When this is done, the station array can be thought of as lying within the boundaries of an airport, e.g., stations #1 and #3 at the ends of runways, station #2 at the intersection of the runways and station #4 in the middle of the field. The altitudes of the three printouts correspond to 200, 1000, and 5000 ft, for this case.

Examination of Fig. 32 shows that, generally, the GDOP in x and y is quite low. Specifically, there are few positions at which the GDOP exceeds three. This simply results from the fact that vehicle positions which are sufficiently far from the station configuration to cause large GDOP in the horizontal plane have not been included in the printouts. Nonetheless, it is concluded that with respect to latitude and longitude, the multilateration technique yields acceptably low GDOP factors when configured as an intercity or approach navigation system.

On the other hand, GDOP factors as high as 500 are experienced in altitude. In particular, at an altitude of two units (200 ft in the approach system) GDOP factors larger than 30 are encountered just outside the immediate vicinity of the station configuration. At an altitude of 10 units, GDOP as high as 100 is encountered and the GDOP factors just outside the station configuration are greater than 10. Finally, at an altitude of 50 units, the maximum GDOP is around 20 and just outside the station configuration it is around two.

In general, then, the GDOP of altitude measurements increases as altitude is decreased and/or distance from the station configuration is increased. Examination of Fig. 32a (with dimensions given in hundreds of feet) shows that great care must be taken in laying out the station configuration if such a technique is to be used for an approach or landing system which provides precise altitude information. This conclusion assumes that the system's ranging measurement error is at least 10 ft rms. Such accuracy is difficult to achieve in two-way systems and very difficult to achieve in one-way ranging systems employing time reference systems.

```
SPECIFY ALTITUDE:
2.0

SPECIFY STATION LOCATIONS:
  40.0,  40.0,180.0,  80.0
(-50.0,  40.0,  40.0,  80.0
   0.0,   0.0,   0.0,   0.2

SPECIFY VARIANCES: SS1,SS2,SS3,SS4,=1,1,1,1

          3.12E+00  2.99E+00  2.89E+00  2.84E+00  2.82E+00  2.86E+00M
+00                                    E                          Y=3:■
          2.77E+02  2.93E+02  3.27E+02  3.76E+02  4.36E+02  5.08E+02

          2.76E+00  2.62E+00  2.52E+00  2.46E+00  2.46E+00  2.51E+00
Y=280       2.40E+00  2.31E+00  2.33E+00  2.45E+00  2.66E+00  2.94E+00
        1.83E+02  1.93E+02  2.26E+02  2.73E+02  3.31E+02  3.99E+02

          2.40E+00  2.25E+00  2.15E+00  2.10E+00  2.10E+00  2.17E+00
Y=240       1.96E+00  1.86E+00  1.89E+00  2.05E+00  2.32E+00  2.65E+00
          1.08E+02  1.12E+02  1.43E+02  1.88E+02  2.42E+02  3.06E+02

          2.06E+00  1.90E+00  1.79E+00  1.73E+00  1.75E+00  1.84E+00
Y=200       1.54E+00  1.42E+00  1.47E+00  1.69E+00  2.02E+00  2.41E+00
          5.74E+01  4.99E+01  7.99E+01  1.22E+02  1.70E+02  2.27E+02

          1.72E+00  1.54E+00  1.44E+00  1.38E+00  1.41E+00  1.54E+00
Y=160       1.17E+00  1.01E+00  1.10E+00  1.39E+00  1.80E+00  2.24E+00
        3.90E+01  7.54E+00  3.97E+01  7.51E+01  1.15E+02  1.62E+02

          1.41E+00  1.19E+00  1.10E+00  1.05E+00  1.10E+00  1.28E+00
Y=120       9.23E-01  7.48E-01  8.81E-01  1.25E+00  1.70E+00  2.17E+00
        4.12E+01  1.47E+01  2.19E+01  4.46E+01  7.10E+01  1.07E+02

          1.23E+00  8.74E-010 8.03E-01  8.02E-01  8.79E-01  1.11E+00
Y= 80       8.88E-01  7.15E-01  8.63E-01  1.26E+00  1.71E+00  2.18E+00
          3.92E+01  1.66E+01  9.99E-01  2.25E+01  3.41E+01  5.82E+01

          1.17E+00  8.66E-01  7.33E-01  7.07E-01  8.17E-01  1.07E+00
Y= 40       1.06E+00  8.66E-01  9.28E-01  1.33E+00  1.80E+00  2.27E+00
          3.48E+01  1.00E+00  1.85E+01  2.28E+01  1.17E+01  1.57E+01

          1.28E+00  1.02E+00  8.56E-01  8.19E-01  9.33E-01  1.17E+00
Y=  0       1.39E+00  1.24E+00  1.30E+00  1.58E+00  1.99E+00  2.44E+00
          5.73E+01  3.66E+01  4.35E+01  5.08E+01  5.03E+01  4.90E+01

           X=000     X=040     X=080     X=120     X=160     X=200
```

Fig. 32a

```
SPECIFY ALTITUDE:
10.0

SPECIFY STATION LOCATIONS:
  40.0,  40.0,180.0,  80.0
(-50.0,  40.0,  40.0,  80.0
   0.0,   0.0,   0.0,   0.2

SPECIFY VARIANCES: SS1,SS2,SS3,SS4,=1,1,1,1

          3.12E+00  3.00E+00  2.93E+00  2.91E+00  2.94E+00  3.02E+00
Y=320       2.96E+00  2.91E+00  2.94E+00  3.07E+00  3.28E+00  3.55E+00
          5.73E+01  6.07E+01  6.83E+01  7.91E+01  9.28E+01  1.09E+02

          2.76E+00  2.62E+00  2.54E+00  2.52E+00  2.55E+00  2.65E+00
Y=280       2.47E+00  2.41E+00  2.46E+00  2.61E+00  2.85E+00  3.16E+00
          3.75E+01  3.98E+01  4.67E+01  5.69E+01  6.96E+01  8.47E+01

          2.41E+00  2.25E+00  2.16E+00  2.14E+00  2.18E+00  2.28E+00
Y=240       2.00E+00  1.92E+00  1.98E+00  2.17E+00  2.46E+00  2.81E+00
          2.21E+01  2.29E+01  2.93E+01  3.88E+01  5.04E+01  6.42E+01

          2.07E+00  1.90E+00  1.79E+00  1.76E+00  1.81E+00  1.94E+00
Y=200       1.55E+00  1.45E+00  1.52E+00  1.77E+00  2.12E+00  2.52E+00
          1.17E+01  1.02E+01  1.63E+01  2.49E+01  3.51E+01  4.72E+01

          1.74E+00  1.55E+00  1.44E+00  1.40E+00  1.46E+00  1.61E+00
Y=160       1.17E+00  1.02E+00  1.13E+00  1.45E+00  1.86E+00  2.31E+00
          7.96E+00  1.86E+00  8.00E+00  1.51E+01  2.33E+01  3.32E+01

          1.45E+00  1.20E+00  1.10E+00  1.07E+00  1.14E+00  1.33E+00
Y=120       9.29E-01  7.57E-01  9.01E-01  1.28E+00  1.72E+00  2.20E+00
          8.36E+00  2.99E+00  4.30E+00  8.78E+00  1.42E+01  2.17E+01

          1.24E+00  9.00E-01  8.13E-01  8.14E-01  8.98E-01  1.14E+00
Y= 80       8.98E-01  7.25E-01  8.74E-01  1.26E+00  1.71E+00  2.19E+00
          7.98E+00  3.33E+00  9.78E-01  4.36E+00  6.80E+00  1.17E+01

          1.20E+00  8.75E-01  7.38E-01  7.14E-01  8.21E-01  1.08E+00
Y= 40       1.07E+00  8.78E-01  9.58E-01  1.35E+00  1.81E+00  2.27E+00
          7.13E+00  9.99E-01  3.69E+00  4.53E+00  2.47E+00  3.35E+00

          1.30E+00  1.03E+00  8.62E-01  8.23E-01  9.31E-01  1.17E+00
Y=  0       1.42E+00  1.27E+00  1.35E+00  1.63E+00  2.03E+00  2.46E+00
          1.18E+01  7.58E+00  8.87E+00  1.03E+01  1.02E+01  9.99E+00

           X=000     X=040     X=080     X=120     X=160     X=200
```

Fig. 32b

Position Errors as a Function of Position for
a Multilateration Altimeter

```
SPECIFY ALTITUDE:
50.0

        3.16E+00  3.04E+00  2.97E+00  2.96E+00  2.99E+00  3.08E+00
Y=320    3.04E+00  3.00E+00  3.04E+00  3.17E+00  3.37E+00  3.64E+00
       1.18E+01  1.25E+01  1.41E+01  1.62E+01  1.90E+01  2.23E+01


        2.80E+00  2.67E+00  2.59E+00  2.57E+00  2.61E+00  2.71E+00
Y=280    2.56E+00  2.51E+00  2.56E+00  2.71E+00  2.95E+00  3.25E+00
       7.88E+00  8.36E+00  9.73E+00  1.18E+01  1.43E+01  1.73E+01


        2.46E+00  2.31E+00  2.22E+00  2.20E+00  2.24E+00  2.35E+00
Y=240    2.11E+00  2.03E+00  2.09E+00  2.28E+00  2.56E+00  2.90E+00
       4.82E+00  5.01E+00  6.25E+00  8.12E+00  1.04E+01  1.32E+01


        2.13E+00  1.96E+00  1.86E+00  1.83E+00  1.88E+00  2.01E+00
Y=200    1.68E+00  1.59E+00  1.67E+00  1.89E+00  2.22E+00  2.62E+00
      2.79E+00  2.55E+00  3.65E+00  5.32E+00  7.34E+00  9.76E+00


        1.82E+00  1.63E+00  1.52E+00  1.49E+00  1.55E+00  1.70E+00
=¥160    1.34E+00  1.22E+00  1.31E+00  1.59E+00  1.98E+00  2.40E+00
      1.97E+00  1.13E+00  1.97E+00  3.32E+00  4.94E+00  6.95E+00


        1.56E+00  1.33E+00  1.22E+00  1.18E+00  1.25E+00  1.44E+00
Y=120    1.12E+00  9.93E-01  1.11E+00  1.43E+00  1.84E+00  2.29E+00
      1.93E+00  9.26E-01  1.12E+00  1.97E+00  3.09E+00  4.63E+00


        1.37E+00  1.11E+00  9.86E-01  9.56E-01  1.04E+00  1.25E+00
=Y80     1.10E+00  9.66E-01  1.09E+00  1.41E+00  1.82E+00  2.27E+00
       1.90E+00  9.31E-01  7.41E-01  1.09E+00  1.63E+00  2.67E+00


        1.32E+00  1.05E+00  9.04E-01  8.69E-01  9.65E-01  1.20E+00
Y= 40    1.25E+00  1.12E+00  1.21E+00  1.52E+00  1.92E+00  2.36E+00
       1.96E+00  9.80E-01  1.02E+00  1.11E+00  9.66E-01  1.32E+00


        1.41E+00  1.15E+00  9.98E-01  9.60E-01  1.06E+00  1.27E+00
Y=  0    1.56E+00  1.45E+00  1.52E+00  1.78E+00  2.14E+00  2.55E+00
      2.93E+00  2.08E+00  2.12E+00  2.30E+00  2.30E+00  2.35E+00


         X=000     X=040     X=080     X=120     X=160     X=200
```

Fig. 32c

Position Errors as a Function of Position of a Multilateration Altimeter

Examination of Fig. 32b (with dimensions given in nautical miles) shows that moderate GDOP is encountered in the vicinity of the station configuration. Therefore, it is concluded that the multilateration altimeter is useful for enroute-intercity vehicles at high (60,000 ft) altitude. Assuming ranging errors on the order of 10-30 ft, maximum altitude error on the order of 200-600 ft will be encountered by vehicles which are less than 50 miles from the nearest station.

SECTION 5

AVAILABILITY OF STABLE SOURCES FOR TIME REFERENCE SYSTEMS

A wide variety of precision frequency standards are available for use in Time Reference Systems. This section describes the characteristics of Hydrogen Maser, Cesium Beam, Rubidium Gas Cell and Quartz Crystal stable sources, the special constraints imposed on them by Airborne Time Reference System requirements, and finally, the applicability of state-of-the-art devices to aircraft communication and navigation problems.

Characteristics of Frequency Standards

Hydrogen maser frequency standards are by far the most stable sources available. In terms of both short-term and long-term frequency stability, the hydrogen maser technique provides at least an order of magnitude improvement relative to any other technique. At the same time, Hydrogen Maser sources are relatively large and consume a relatively large amount of power (800 lbs, 200 watts).

Cesium beam devices are also extremely stable on a long-term basis. Stable sources using cesium beam resonators are of moderate size and consume moderate (65 lbs, 60 watts) power. The short-term stability of such devices is determined by the characteristics of a quartz crystal oscillator which is typically part of the cesium beam frequency standard realization.

The extremely short-term stability of a rubidium standard is similarly determined by the characteristics of a quartz oscillator which is part of the system. The intermediate-term stability exceeds that of cesium sources. But, the long-term stability is inferior to that obtained with a cesium beam. Rubidium standards are lighter and consume less power (40 lbs, 40 watts) than cesium or, of course, hydrogen maser sources.

Figure 33 reviews the fractional frequency deviation as a function of averaging time typically obtainable using the various techniques. First, note that the hydrogen maser source out-performs the others. Note also that at very short averaging times the cesium and rubidium source stabilities are determined by their internal quartz crystal oscillators, as noted above. Finally, note that the quartz and rubidium data are "detrended". This means that the effect of long-term frequency drift (e. g., aging) has been removed from the data before analysis of stability.

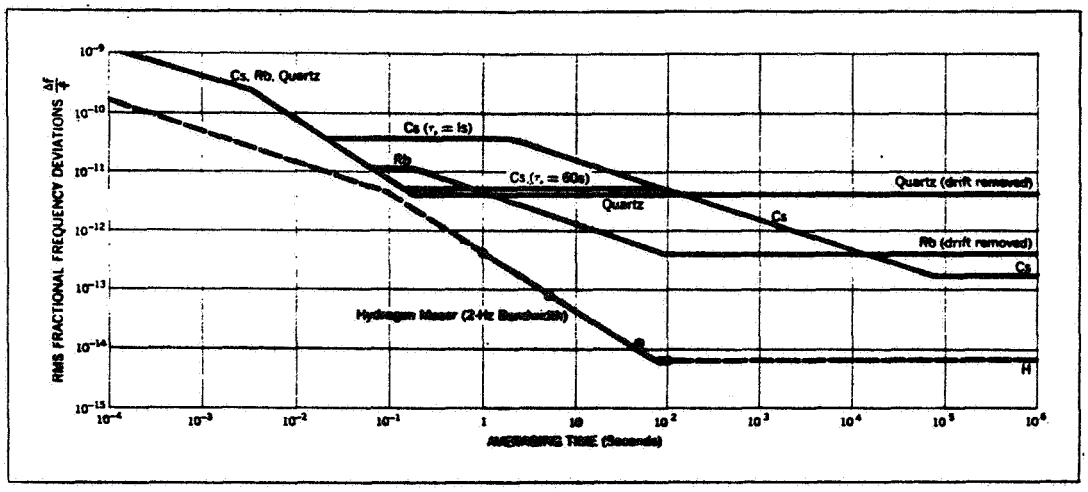


Fig.33 Frequency Standard Stability*

Characterization of Frequency Standards for Time Reference Systems

A time reference system consists of a very stable frequency standard of the type described above and an event counter. The contents of the counter is proportional to time. Therefore, the accuracy of the time reference is derivable from the frequency stability specification.

Time reference systems typically impose severe constraints on the frequency standard specifications. This comes about because the time reference output tends to accumulate the effects of the frequency fluctuations. Consider, for example, an oscillator whose long-term average frequency is exactly known. At any instant of time the oscillator output may be offset in frequency by, say, Δf Hz. For the situation under consideration here, the average value of Δf, $\langle \Delta f \rangle$, is zero. However, the average value of the incremental phase accumulated due to the fluctuations, $\Delta f(t)$, is not zero. In fact, the phase exhibits a random walk away from zero and after an infinite time it will have an infinite rms error. This means that a time reference system which is driven by a stable source having no long-term error but only mean-zero short-term fluctuations will have a continually increasing rms timing error.

The rms accumulated phase (or timing error) as a function of averaging time is directly derivable† from the fractional frequency deviation as a function of averaging time such as is given in Fig.33. In particular, the

* From the HP 1969 catalog. The stabilities shown are not necessarily state-of-the-art, but, they show the relative merits of the various techniques.

† Cutler and Searle, Frequency Fluctuations in Frequency Standards, Proc. IEEE, Feb., 1966, p.139.

variance of the fractional frequency deviation is given by

$$\sigma^2 \frac{[\langle \dot{\phi} \rangle]}{\omega_o^2} = \frac{2}{(\tau \omega_o)^2} [R_\phi(0) - R_\phi(\tau)] \tag{66}$$

where τ is the averaging time, ω_o is the oscillator center frequency, and $R_\phi(\tau)$ is the autocorrelation of the phase fluctuations of the oscillator referenced to ω_o. The normalized variance of the accumulated phase error is given by

$$\sigma^2 \frac{[\Delta_\tau \phi]}{\omega_o^2} = \frac{2}{\omega_o^2} [R_\phi(0) - R_\phi(\tau)] \tag{67}$$

Thus, the rms accumulated phase error in terms of the fractional frequency deviation is, from Equations (66) and (67)

$$\sigma \frac{[\Delta_\tau \phi]}{\omega_o} = \tau \sigma \frac{[\langle \dot{\phi} \rangle]}{\omega_o} \tag{68}$$

This relationship is used to derive Fig.34 which shows the accumulated phase errors referenced to 1 GHz expected from various state-of-the-art oscillators. Again, as in the case of Fig.33, the effect of frequency drift has been removed from the quartz oscillator and the rubidium reference data.

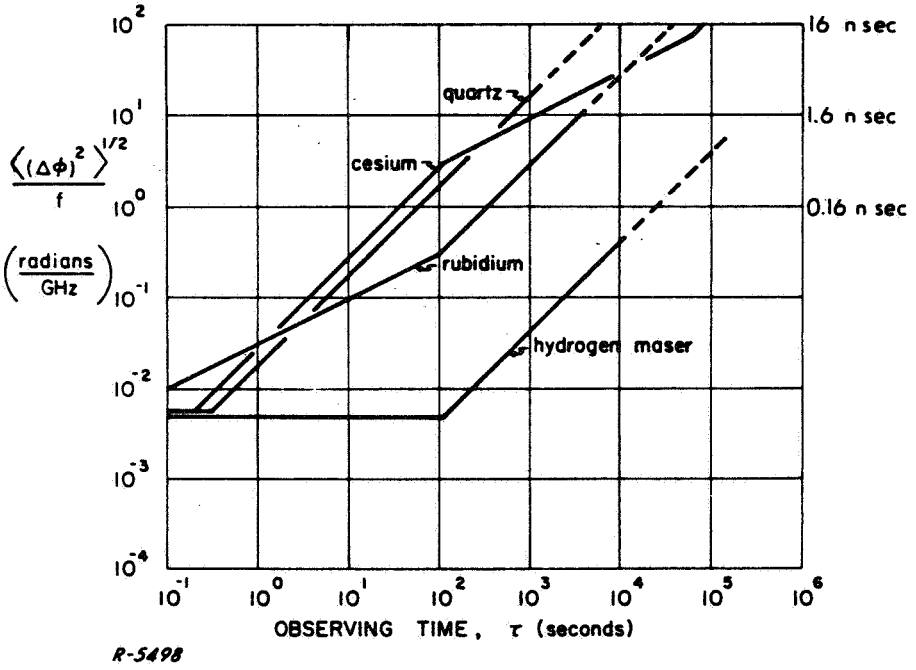


Fig.34 Normalized Phase Departure vs. Observation Time for Various Frequency Standards. Dotted Lines are Extrapolations.

Let us now consider the effect of a long-term drift on timing reference system error. For Figures 33 and 34, the uncertainty in the oscillator center frequency, ω_o, was considered to be unmeasurable in the case of hydrogen maser or cesium beam standards or the effect of this uncertainty was removed in the case of rubidium and quartz standards.

When the center frequency ω_o is in error by $\Delta\omega_o$ due to a calibration error or an uncompensated drift, the input to the event counter portion of the TRS is of the form

$$e(t) = \sin(\omega_o + \Delta\omega_o)t \tag{69}$$

Thus, zero crossings of e(t) occur at

$$t_{z.c.}(\Delta\omega) = \frac{n\pi}{\omega_o + \Delta\omega_o} \qquad n = 0, 1, 2 \ldots \infty \tag{70}$$

These should occur at

$$t_{z.c.}(0) = \frac{n\pi}{\omega_o} \qquad n = 0, 1, 2 \ldots \infty \tag{71}$$

The error due to a center frequency uncertainty $\Delta\omega_o$ is therefore

$$\Delta t = t_{z.c.}(\Delta\omega) - t_{z.c.}(0) = \frac{-n\pi\Delta\omega_o}{\omega_o(\omega_o + \Delta\omega_o)} \qquad n = 1, 2, \ldots \infty \tag{72}$$

and normalizing by dividing Equation 7 by Equation 6 yields the fractional timing error

$$\frac{\Delta t}{t_{z.c.}} = -\frac{\Delta\omega_o}{\omega_o + \Delta\omega_o} \approx -\frac{\Delta\omega_o}{\omega_o} \tag{73}$$

which, it is noted, has the same form as Equation 68.

The timing error introduced by calibration error or uncompensated drift can be large. For example, if an uncertainty of center frequency of 1 part in 10^9 exists, then the timing error accumulates at a rate of 1 nsec per second.

Applicability to Passive Navigation Systems

The discussion of Section 4 indicates that effective utilization of time reference systems for passive navigation requires that the absolute timing error of the on-board system be less than 50-100 nsec. The curve of

Figure 34 for a quartz oscillator shows that a timing error accumulation of 0.1 nanosecond in 10 seconds can be expected from a state-of-the-art source in a laboratory environment provided the drift error is removed or compensated. This figure corresponds almost exactly with the rms phase deviation specification given for the HP model 105A,B oscillator. For the sake of example, let us assume that the total frequency uncertainty due to drift, calibration, and variations in temperature, humidity, supply voltage and load is on the order of 1×10^{-11}, then, this source of error would also contribute 0.1 nsec in 10 seconds to the accumulated timing error. It is therefore optimistic to expect a total rms timing error of, say, 0.2 nanosecond/10 seconds, or 100 nanoseconds in 5000 seconds. The error from such an oscillator would, thus, exceed requirements after less than 1-1/2 hrs of operation assuming perfect synchronization at the start of the period.

The frequency uncertainty of 1×10^{-11} assumed in the above discussion would be rather difficult to obtain even in a controlled environment. Somewhat worse results using quartz oscillators should be expected in the airborne environment. Two specially equipped cesium beam standards were flown around the world recently.* The clocks produced more than an order of magnitude improvement in performance over that estimated above. Specifically, the accumulated timing error per day was on the order of 100 nanoseconds. Such sources cost roughly $15,000 which is an order of magnitude larger than the cost of the precision quartz oscillator assumed for the error estimates above.

---

* HP flying clock experiments. See the HP Journal, December, 1967.

APPENDIX A

MULTIPATH

At times, the operation and performance of air-to-ground and ground-to-air communication and navigation links are severly degraded due to the presence of multipath. The anomalous behavior of ILS and VOR equipments caused by the poor electromagnetic environment are of current concern because several accidents are attributable to such causes. This section is devoted to identifying sources of multipath disturbance and to analyzing the effect of multipath on a variety of navigation equipments with specific emphasis on VOR/VORTAC systems.

Siting Errors

Azimuth and distance measurement errors normally labeled as "siting errors" are usually attributed to features of the surroundings of the station which can introduce anomalous propagation characteristics into the signal received by the airborne equipment. Such anomalous effects include multipath transmission, through both specular-like and diffuse reflections, scattering of the electromagnetic energy by rough elevated surface features, and apparent counterpoise imbalance due to significant nonuniformities in the electrical characteristics of the surrounding terrain.

In the discussion below, a general survey of reflectors is considered for the cases of single and multiple reflectors, and for surfaces that depart from the model of a smooth, uniform earth because of roughness and non-homogeneity of surface dielectric constants. Then, the requirements of the practical site with emphasis on artificial reflectors prevalent at VOR facilities are considered. The production of scalloping errors by nearby flying aircraft is also considered.

A. General Survey of Reflectors and Scatters

1. Single Reflectors

These reflectors can be of the single tower or tree type, or a larger structure such as a hangar face or mountain face. The latter case is distinguished by being large electrically in all of its dimensions. For both types of structure an analytic method equivalent to ray tracing in optics is suitable. That is to say, a single strong reflection is present at any given space angle heading which is equivalent to a multipath on the original signal. The analytic model follows from this point of view.

2. Rough Elevated Surface Effects

This classification includes several different effects due to large-rough-edged scatters. Shadowing of the aircraft can occur for very large objects in the flight path. Mountain top siting and the use of doppler VOR can alleviate this condition in some cases but it remains a difficult problem in rugged, mountainous terrain. Knife-edge diffraction can occur which scatters energy toward strong reflectors not in the path of the line-of-sight between aircraft and VOR station. Multipath from these reflectors can cause scalloping. Lastly, diffuse reflections due to surface or profile roughness can take place from a hillside or mountainside, simultaneously with a strong specular reflection.

Diffuse reflection is a phenomenon also associated with seasonal foliage increase on wooded slopes. Reflection can change from mainly specular to mainly diffuse between winter and summer. In addition to foliage variation, other terrain variations are caused by short-term and seasonal weather patterns which affect soil dielectric constant and reflectivity. Long-term changes due to natural accretion of foliage must also be considered in terms of site maintenance.

3. Non-Uniformity of Electrical Surface Properties

These are of two main types: land-water interfaces, and diffraction (scattering) by the structure of an otherwise fairly level ground, in particular, roughness and horizontal variations of dielectric constant. The effect of a land-water interface is to provide a very strong single multipath signal which can produce severe nulls in the vertical plane. This effect is observed experimentally at VOR installations near such interfaces. When a transmitter is set up near a coast or on an island, the reflection pattern may undergo variations related to the tides, which makes calibration difficult. Again, the use of doppler VOR can reduce the deep nulls caused by sea water reflection.

In order to estimate how closely the ground approaches an ideal reflecting surface, Rayleigh's criterion for roughness is usually invoked. Originally developed for optical purposes, Rayleigh's criterion has been applied with good success to VHF propagation problems and its use is well known. Figure A-1 shows how this criterion is applied to setting specifications for terrain near a VOR station. In essence the criterion sets the limit on the height of objects for which the specular component far exceeds the diffuse.

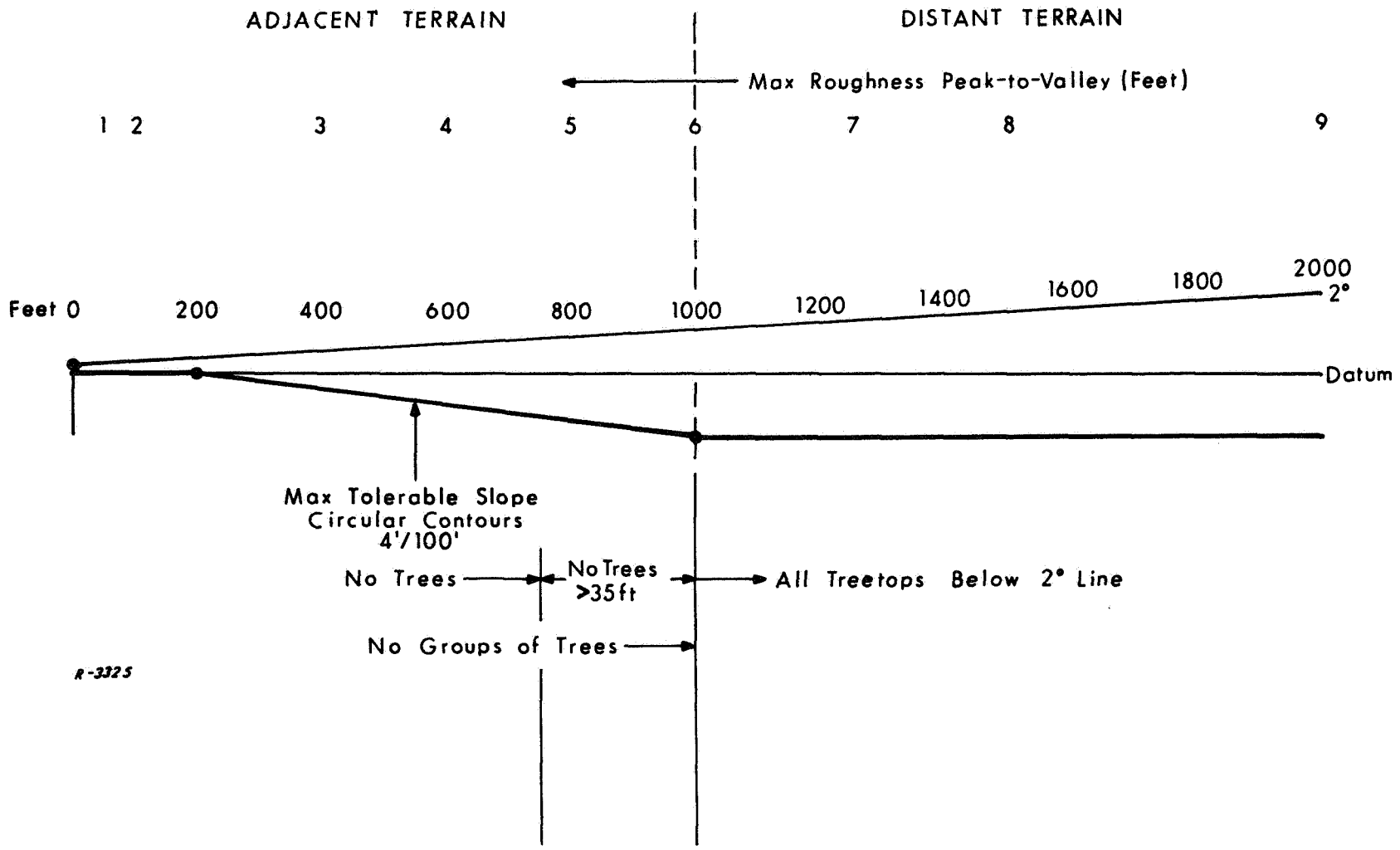


Fig. A-1 Practical Site Requirements for Satisfactory Performance Natural Terrain Criteria

Horizontal variations in dielectric constant can occur at terrain interfaces, e.g., woods-grass land. At VHF frequencies and for horizontal polarization, the reflection coefficient of most soil types may be taken as unity with a 180^{o} phase reversal; however, a more careful treatment must be given at land-water interfaces because of the substantial differences in dielectric constant.

Seasonal changes in the ratio of diffuse to specular components of the scattering referred to above for the case of wooded slopes are also applicable here to the case of horizontal terrain where foliage changes may contribute significantly to variations in scalloping error. This latter effect has also been observed experimentally.*

B. Man-made Site Features

At terminal VOR stations, man-made reflectors constitute the worst offenders as far as scalloping is concerned. Due to the construction of runways, terrain is usually cleared and leveled for thousands of feet in the vicinity of airports, although in coastal regions land-water interfaces are common at TVOR sites.

However, it is not possible to control distant terrain effects near airports because of their proximity to built-up metropolitan areas.

A partial listing of troublesome man-made reflectors is as follows:

1) metal frame or reinforced concrete buildings,
2) metal fences,
3) power lines,
4) towers,
5) hangars,
6) bridges and elevated highways, and
7) grain elevators and storage tanks.

---

* L. H. W. Feyer and K. Nattrodt, "Test Results from a Tower VOR Over a Forest Area," IEEE Transactions on Aerospace and Navigational Electronics, March 1965, No. 1.

For each of the reflectors listed, some general comments can be made. The effect of these reflectors on scalloping error is always proportional to their size, proximity to the VOR installation, height with respect to the VOR antenna, and reflection coefficient. The reflectivity is large since all objects listed are largely metallic.

The effect of reflections from long, thin conductors such as power lines and fences is largely dependent on the orientation of the reflectors. Fences and power lines that lie substantially along radials produce minimum reflections and can be used within a few hundred feet of the antenna.

Effects due to high towers are well known and can result in troublesome scalloping. Similarly, poor results are obtained in the vicinity of tall grain elevators and fuel storage tanks. Bridges have long horizontal reflecting surfaces and in some cases their height is significant.

C. Moving Single Reflectors

One type of reflector which has not been discussed in the past when considering siting criteria is moving aircraft in the vicinity of the VOR station. The effect of other aircraft on flight tests is clearly evident. The other aircraft act as moving reflectors and it is not difficult to imagine a geometry in which the reflecting aircraft can give fairly constant scalloping. The reflection coefficient of aircraft can be obtained from controlled flight tests, and by adapting results of work on radar cross sections.* As an example of the size of disturbance, a bearing error of more than a degree was observed at Doppler Station No. 2 at NAFEC, Atlantic City, when an aircraft was flying over the station.

Calculation of Azimuth Error in VOR, DVOR and PDVOR Caused by One Reflector

We shall now consider that case of two-path propagation--a direct path and a reflected path--and compute the change in the azimuth reading that the presence of one extraneous path will cause in VOR and DVOR.

---

* Crispin, Jr., J.W. and Maffett, A.L., "Radar Cross Section Estimation for Complex Shapes," Proc. IEEE, Vol. 53, No. 8, pp. 972-982, August 1965.

The purpose of this computation is to illustrate the basic manipulations involved in analyzing multipath effects on navigation-aid systems, and examine the results to establish a basis for correcting for the effects of one reflector by introducing another suitably positioned reflector. An important by-product of this computation is basic formulas for comparing VOR, DVOR and PDVOR sensitivities to two-path propagation.

Direct Plus Specular-Reflected Path: Conventional VOR

The conventional VOR signal consists of a modulated RF carrier containing the following modulations:

(a) A phase-reference tone at $\rho = 2\pi \times 30$ rad/sec, frequency-modulated on a subcarrier at $\omega_{sc}/2\pi = 9960$ Hz, yielding $\cos [\omega_{sc} t + \delta \sin \rho t]$, with $\delta = 16$ rad. This result appears, in turn, as 30% amplitude modulation on the RF carrier.

(b) A tone at $\rho/2\pi = 30$ Hz, described by $\sin (\rho t + \theta_s)$, where θ_s is a phase angle that measures the azimuth of the airborne receiver relative to the station reference. This tone appears as 30% amplitude modulation on the RF carrier.

(c) Intelligence consisting of voice or a coded tone serving as a station identification signal, amplitude modulated on the RF carrier.

Of main concern in this discussion is the effect of two-path propagation on the measurement of the azimuth angle, θ_s, in the VOR receiver. Accordingly, we express the VOR signal in the following alternate forms:

$$e_{VOR}(t) = E_s \{1 + m_1 \sin (\rho t + \theta_s) + m_2 \cos [\omega_{sc} t + \delta \sin \rho t]\} \cos [\omega_c t + \phi_c] \qquad (A1)$$

$$= E_s \cos [\omega_c t + \phi_c]$$

$$+ m_1 E_s \sin (\rho t + \theta_s) \cos [\omega_c t + \phi_c]$$

$$+ m_2 E_s \cos [\omega_{sc} t + \delta \sin \rho t] \cos [\omega_c t + \phi_c] \qquad \text{(A2)}$$

$$= E_s \cos (\omega_c t + \phi_c)$$

$$+ \frac{1}{2} m_1 E_s \sin [(\omega_c + \rho) t + \theta_s + \phi_c]$$

$$+ \frac{1}{2} m_1 E_s \sin [(\omega_c - \rho) t - \theta_s + \phi_c]$$

$$+ \frac{1}{2} m_2 E_s \cos [(\omega_c + \omega_{sc}) t + \phi_c + \delta \sin \rho t]$$

$$+ \frac{1}{2} m_2 E_s \cos [(\omega_c - \omega_{sc}) t + \phi_c - \delta \sin \rho t] \qquad \text{(A3)}$$

where m_1 and m_2 are the amplitude modulation indices (= 30%) and $\cos (\omega_c t + \phi_c)$ represents the RF carrier.

The signal received by a distant aircraft in general consists of the direct signal plus reflections from a number of terrain features. In this analysis we consider only a single specular reflector to illustrate the method of computing the azimuth error introduced by multipath. This formulation, however, can be extended to the more general case of multiple reflectors, including scattering.

The single-reflector station model is illustrated in the sketch of Fig. A-2. With the direct-path signal expressed in the second form given above, the assumed specularly reflected signal is given by

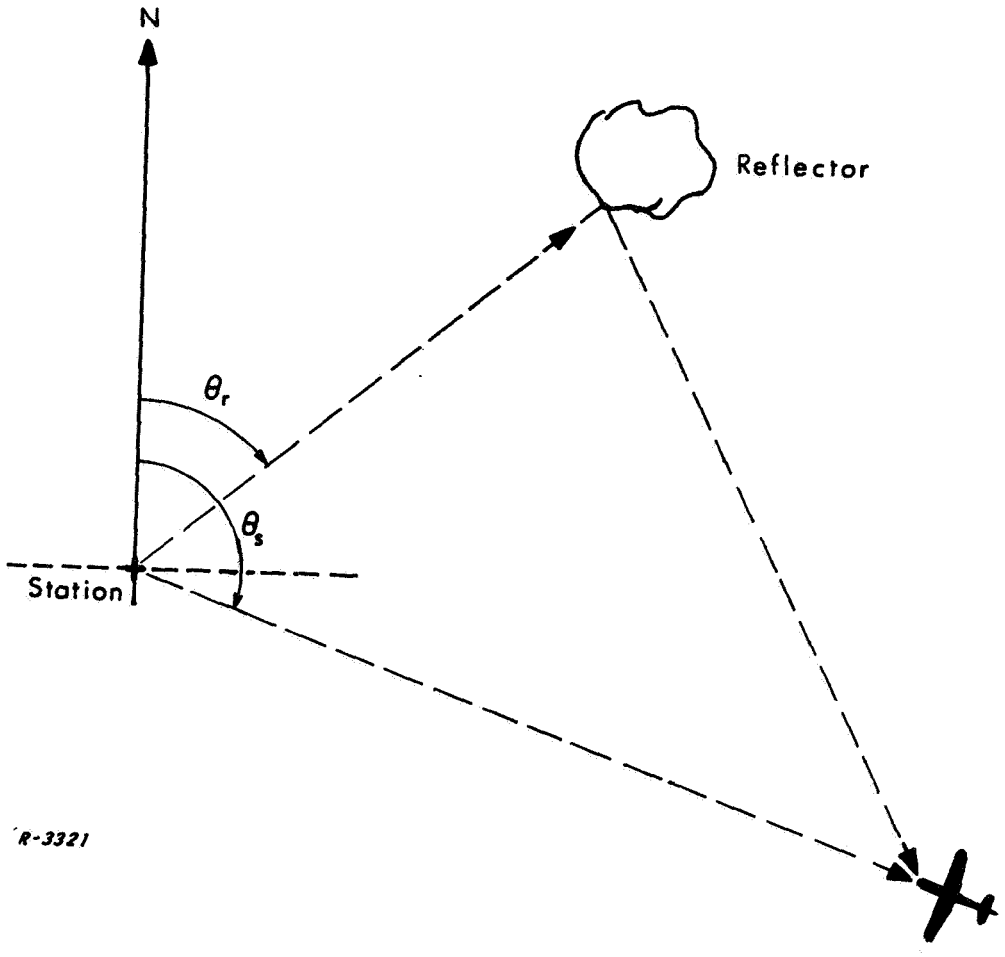


Fig. A-2 Geometrical Configuration of Station, Reflector and Aircraft

$$e_r(t) = aE_s \cos[\omega_c(t-t_d) + \phi_c]$$

$$+ am_1 E_s \sin[\rho(t-t_d) + \theta_r] \cos[\omega_c(t-t_d) + \phi_c]$$

$$+ am_2 E_s \cos[\omega_{sc}(t-t_d) + \delta \sin \rho(t-t_d)] \cos[\omega_c(t-t_d) + \phi_c] \quad \text{(A4)}$$

where $\underline{a}$ represents the ratio of reflected to direct signal, and t_d the delay difference between the two signal replicas. Accordingly, $\underline{a}$ is a function of the reflection properties of the reflector, the position of the reflector relative to the station as well as to the airborne receiver. The direct signal is usually much stronger than the sum of all reflected signals as evidenced by the fact that bearing errors due to the reflections, even in the worst case, seldom exceed several degrees. However, we will not make any assumptions concerning $\underline{a}$ until we derive a general result, after which we will consider the special case of small $\underline{a}$.

The carrier resultant of the direct and reflected signal is given by

$$e_c(t) = E_s A_c(a, t_d) \cos[\omega_c t + \phi_c + \theta_c(a, t_d)] \quad \text{(A5)}$$

where

$$A_c(a, t_d) = \sqrt{1 + a^2 + 2a \cos \omega_c t_d} \quad \text{(A6)}$$

$$\theta_c(a, t_d) = -\tan^{-1} \frac{a \sin \omega_c t_d}{1 + a \cos \omega_c t_d} \quad \text{(A7)}$$

If we let

$$\phi_{c1} = \phi_c - \tan^{-1} \frac{a \sin \omega_c t_d}{1 + a \cos \omega_c t_d} \quad \text{(A8)}$$

then, for the AM demodulator in the receiver, the carrier reference is

$$\cos\left[\omega_c t + \phi_{c1}\right] \tag{A9}$$

AM Tone Detection

Envelope detection followed by lowpass filtering yields the AM, variable-phase tone

$$E_{AM}(t) = k_d m_1 E_s \cos(\phi_{c1} - \phi_c)\left[\sin(\rho t + \theta_s) + a' \sin(\rho t + \theta_r - \rho t_d)\right] \tag{A10}$$

where

$$\cos(\phi_{c1} - \phi_c) = \frac{1 + a\cos\omega_c t_d}{\sqrt{1 + a^2 + 2a\cos\omega_c t_d}} \tag{A11}$$

$$a' \equiv a\,\frac{\cos(\omega_c t_d + \phi_{c1} - \phi_c)}{\cos(\phi_{c1} - \phi_c)} \tag{A12}$$

and k_d is a proportionality constant characteristic of the detector. The detected AM tone may be expressed in resultant form as

$$E_{AM}(t) = k_d m_1 E_s A_{AM}(a, t_d) \sin(\rho t + \phi_{\epsilon, AM}) \tag{A13}$$

where $A_{AM}(a, t_d)$ and $\phi_{\epsilon, AM}$ are as defined in Fig. A-3. The resultant phase error on the information bearing tone is, therefore,

$$\phi_{\epsilon, AM} = -\tan^{-1}\frac{a' \sin(\rho t_d + \theta_s - \theta_r)}{1 + a' \cos(\rho t_d + \theta_s - \theta_r)} \quad \text{rad} \tag{A14}$$

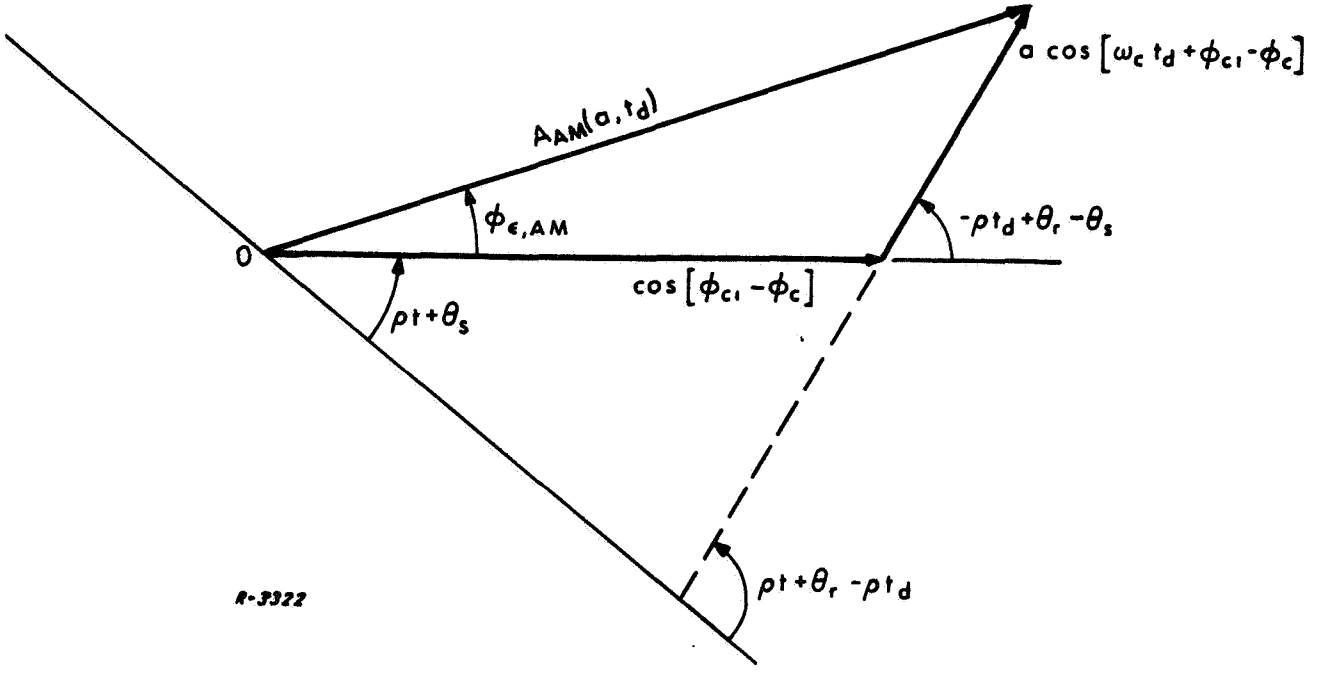


Fig. A-3 Phasor Diagram Representation of the Sum of Direct and Reflected Path Contributions to the AM Tone in Conventional VOR

Note that since $2\pi/\rho = 1/30$ sec, $\rho t_d \leq 10^{-2}$ rad for all $t_d \lesssim 0.1$ msec, and we can write

$$\phi_{\epsilon, AM} \approx -\tan^{-1} \frac{a' \sin(\theta_s - \theta_r)}{1 + a' \cos(\theta_s - \theta_r)} \quad \text{for } |\theta_s - \theta_r| \geq 0.1 \text{ rad} \qquad \text{(A-15)}$$

The significance of the limitation on t_d is best appreciated by recalling that 0.1 msec of delay difference between direct and reflected paths corresponds to 30 kilometers of path-length difference. Since such a path-length difference is generally infrequent, the above approximation for $\phi_{\epsilon, AM}$ is generally reasonable.

A brief examination of a' is also in order. We first observe that

$$a' = a\left[\cos \omega_c t_d - \sin \omega_c t_d \tan (\phi_{c1} - \phi_c)\right] \tag{A16}$$

$$= a\left[\cos \omega_c t_d + \frac{a \sin^2 \omega_c t_d}{1 + a \cos \omega_c t_d}\right] \tag{A17}$$

$$= a\,\frac{a + \cos \omega_c t_d}{1 + a \cos \omega_c t_d} \tag{A18}$$

$$= 1 - \frac{1 - a^2}{1 + a \cos \omega_c t_d} \tag{A19}$$

The expressions in Equations (A18) and A19) show that $|a'| \leq a$ for all values of $\omega_c t_d$, and indeed that $-a \leq a' \leq a$ for $-1 \leq \cos \omega_c t_d \leq 1$. For $\omega_c t_d = \cos^{-1}(-a)$, $a' = 0$. A plot of (a'/a) vs $\omega_c t_d$ for various values of $\underline{a}$ is given in Fig. A-4.

Thus, for $a \leq 0.1$, we can write

$$\phi_{\epsilon, AM} \approx -a' \sin (\rho t_d + \theta_s - \theta_r), \quad |a| \leq 0.1 \tag{A20}$$

$$\approx -a' \sin (\theta_s - \theta_r), \quad |a| \leq 0.1, \quad t_d \leq 10^{-4} \text{ sec}$$

$$\text{and } |\theta_s - \theta_r| \geq 10^{-1} \text{ rad} \tag{A21}$$

It is interesting to observe that the variation of a' with $\omega_c t_d$, illustrated in Fig. A-4, predominately determines the fluctuations in $\phi_{\epsilon, AM}$ when the flight course is such that $(\theta_s - \theta_r)$ remains nearly constant or changes slowly, as for a radial approach to the station.

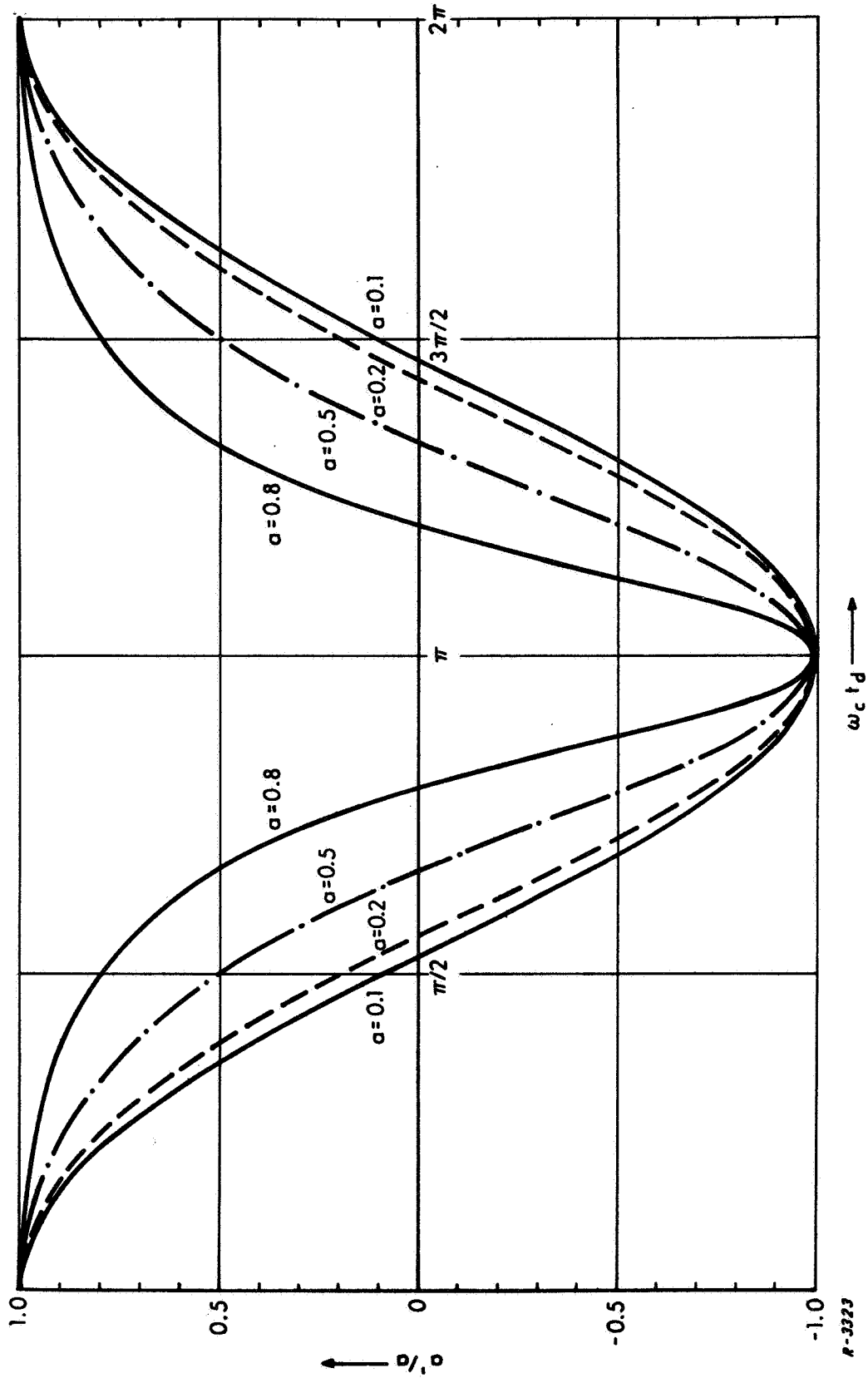


Fig. A-4 Variation of a'/a with $\omega_c t_d$ for Various Values of a

FM Subcarrier Detection

The output of the receiver envelope detector also yields the frequency modulated subcarrier that carries the phase-reference tone. The detected FM subcarrier is given by

$$E_{FM}(t) = k_d m_2 E_s \cos(\phi_{c1} - \phi_c) \left[\cos[\omega_{sc} t + \delta \sin \rho t] + a' \cos\{\omega_{sc}(t-t_d) + \delta \sin[\rho(t-t_d)]\}\right] \quad \text{(A22)}$$

where a' and $\cos(\phi_{c1} - \phi_c)$ are defined in Equations (A11) and (A12). The detected FM subcarrier can be expressed in resultant form as

$$E_{FM}(t) = k_d m_2 E_s A_{FM}(t) \cos[\omega_{sc} t + \delta \sin \rho t + \theta(t)] \quad \text{(A23)}$$

where

$$A_{FM}(t) = \cos(\phi_{c1} - \phi_c) \sqrt{1 + a'^2 + 2a' \cos \phi_F(t)} \quad \text{(A24)}$$

$$\theta(t) = -\tan^{-1} \frac{a' \sin \phi_F(t)}{1 + a' \cos \phi_F(t)} \quad \text{(A25)}$$

$$= \mathrm{Im}\left\{-\sum_{n=1}^{\infty} \frac{(-a')^n}{n} e^{jn\phi_F(t)}\right\}, \; a'^2 < 1 \quad \text{(A26)}$$

and

$$\phi_F(t) = \omega_{sc} t_d + \delta \sin \rho t - \delta \sin[\rho(t-t_d)] \quad \text{(A27)}$$

We shall now derive an expression for the reference phase error $\phi_{\epsilon, FM}$ added to θ_s by the presence of the reflected signal. To this end, we first express Equation (A27) in a slightly different form, as

$$\phi_F(t) = \omega_{sc} t_d + \delta_d \sin(\rho t + \phi_d) \tag{A28}$$

where

$$\delta_d = 2\delta \sin\left(\frac{\rho t_d}{2}\right) \tag{A29}$$

$$\phi_d = -\frac{(\rho t_d + \pi)}{2} \tag{A30}$$

Substitution in Equation (A25) followed by use of the identity

$$e^{jx \sin \beta} = \sum_{k=-\infty}^{\infty} J_k(x)\, e^{jk\beta}$$

where $J_k(x)$ is the Bessel function of the first kind, k^{th} order and argument x, leads to

$$\theta(t) = -\sum_{k=-\infty}^{\infty} \sum_{n=1}^{\infty} \frac{(-a')^n}{n} J_k(n\delta_d) \sin\left[k\rho t + k\phi_d + n\omega_{sc} t_d\right], \quad a'^2 < 1 \tag{A31}$$

The FM subcarrier in the output of the envelope detector is fed into an amplitude insensitive discriminator which ideally delivers a voltage directly proportional to the composite frequency fluctuation in Equation (A23). We will assume that negligible distortion is introduced by the discriminator so that the output waveform is a good reproduction of the frequency fluctuation. The output of the discriminator is thus, with $\Delta\Omega = \rho\,\delta$,

$$e_{disc}(t) \propto \Delta\Omega \cos \rho t + \dot{\theta}(t) \tag{A32}$$

$$= \Delta\Omega \cos \rho t$$

$$- \sum_{k=-\infty}^{\infty} \sum_{n=1}^{\infty} k\rho \frac{(-a')^n}{n} J_k(n\delta_d) \cos \left[k\rho t + k\phi_c + n\omega_{sc} t_d\right], \; a'^2 < 1$$

$$= \Delta\Omega \cos \rho t + \sum_{k=1}^{\infty} D_k(t) \tag{A33}$$

where

$$D_k(t) = -2k\rho \left[\sum_{n=1}^{\infty} \frac{(-a')^n}{n} J_k(n\delta_d) \cos (n\omega_{sc} t_d)\right] \cos k(\rho t + \phi_d)$$

$$\text{for } k = 1, 3, 5, \ldots \tag{A34}$$

$$= 2k\rho \left[\sum_{n=1}^{\infty} \frac{(-a')^n}{n} J_k(n\delta_d) \sin (n\omega_{sc} t_d)\right] \sin k(\rho t + \phi_d)$$

$$\text{for } k = 2, 4, 6, \ldots \tag{A35}$$

Equations (A32) and (A33) clearly show that the presence of the second path gives rise to phase or frequency modulation distortion terms at the frequency of the modulating tone, ρ, and at all harmonics of it. Of particular interest here is the distortion term that coincides in frequency with the tone at ρ rad/sec. This term is given by

$$D_1(t) = -\left[2\rho \sum_{n=1}^{\infty} A_{1,n}(a', \delta, \omega_{sc} t_d)\right] \cos (\rho t + \phi_d) \tag{A36}$$

and

$$A_{1,n}(a', \delta, \omega_{sc} t_d) \equiv \frac{(-a')^n}{n} J_1(n\delta_d) \cos(n\omega_{sc} t_d) \tag{A37}$$

This fundamental distortion term is seen to differ in phase from the phase of the bearing signal by $\phi_d - \pi = -[\rho t_d + \pi]/2$. The resultant reference tone will therefore have a phase error given by

$$\phi_{\epsilon, FM} = -\tan^{-1} \frac{\gamma_1 \cos(\rho t_d/2)}{1 - \gamma_1 \sin(\rho t_d/2)} \tag{A38}$$

where

$$\gamma_1 \equiv (2/\delta) \sum_{n=1}^{\infty} A_{1,n}(a', \delta, \omega_{sc} t_d), \quad a'^2 < 1. \tag{A39}$$

The infinite series in Equation (A39) as well as those for the evaluation of the harmonic distortion components in Equations (A34) and (A35) converge rather slowly for values of a' nearing unity. However, since we are mainly interested in establishing error bounds, it is interesting to observe that

$$|A_{1,n}(a', \delta, \omega_{sc} t_d)|_{max} < 0.6 \frac{a'^n}{n}$$

where we made use of the property $|J_1(\delta)|_{max} < 0.6$. The value of $A_{k,n}$ for the k^{th} harmonic may be similarly bounded and thus

$$|\gamma_1|_{max} < \frac{1.2}{\delta} \sum_{n=1}^{\infty} \frac{a'^n}{n} = -\frac{1.2}{\delta} \ell n(1 - a'), \quad a'^2 < 1 \tag{A40}$$

Therefore, from Equation (A38) we have

$$|\phi_{\epsilon, FM}| < \left| \tan^{-1} \frac{\cos(\rho t_d/2)}{\delta/[1.2 \ell n(1 - a')] + \sin(\rho t_d/2)} \right| \tag{A41}$$

In view of the fact that negative and positive quantities are summed in the expression for γ_1 and only the magnitudes of these quantities are considered in Equation (A40), the upper bounds in Equations (A40) and (A41) are not expected to be sufficiently tight to make them useful in obtaining realistic estimates of $|\gamma_1|_{max}$ and of the maximum phase error. But for values of the parameters involved that yield small values for the bounds, it is useful to know that even the most pessimistic estimates of $|\gamma_1|_{max}$ and of the phase error would be that small.

Special Case of Small a

The above results may be greatly simplified in the special case in which the amplitude ratio a is less than 0.1. First we observe that

$$\theta(t) = \tan^{-1} \frac{a' \sin \phi(t)}{1 + a' \cos \phi(t)} = a' \sin \phi(t) - \frac{1}{2} a'^2 \sin 2\phi(t) + \frac{1}{3} a'^3 \sin 3\phi(t) - \ldots$$

$$\approx a' \sin \phi(t) \text{ for all } a' \leq 0.2 \tag{A42}$$

From Equation (42) we have

$$\theta(t) = -\dot{\phi}(t) \sum_{n=1}^{\infty} (-a')^n \cos n\phi(t)$$

$$\approx a' \ddot{\phi}(t) \cos \phi(t) \text{ if } a' \leq 1/10 \tag{A43}$$

Under these conditions, an ideal carrier frequency demodulator would deliver an output, prior to subchannel filtering, proportional to

$$\dot{\theta}(t) \approx \Delta\Omega \cos \rho t + a' \dot{\phi}(t) \cos \phi(t), \quad a' \leq 1/10 \tag{A44}$$

By direct substitution and expansion of Equation (A44), or direct use of Equations (A34) and (A35) with n restricted only to unity, as per approximation (A43), we have, for $\underline{a}' \leq 1/10$,

$$D_k(t) \approx \left[2\, a'\, \rho\, k J_k(\delta_d) \cos \omega_{sc} t_d\right] \cos k(\rho t + \phi_d)$$

$$\text{for } k = 1, 3, 5, \ldots \qquad \text{(A45)}$$

$$\approx \left[2\, a' \rho k J_k(\delta_d) \sin \omega_{sc} t_d\right] \sin k(\rho t + \phi_d)$$

$$\text{for } k = 2, 4, 6, \ldots \qquad \text{(A46)}$$

Moreover, in Equation (A39) for the phase error in the resultant fundamental term,

$$\gamma_1 \approx -(2a'/\delta)\, J_1(\delta_d) \cos \omega_{sc} t_d \qquad \text{(A47)}$$

which upon substitution in Equation (A38) yields

$$\phi_{\epsilon, FM} \approx \tan^{-1} \frac{(2a'/\delta)\, J_1(\delta_d) \cos(\omega_{sc} t_d) \cos(\rho t_d/2)}{1 + (2a'/\delta)\, J_1(\delta_d) \cos(\omega_{sc} t_d) \sin(\rho t_d/2)} \qquad \text{(A48)}$$

For VOR systems, δ is a fixed parameter equal to 16. Since $|J_1(\delta_d)| < 0.6$, and $|a'| \leq a \leq 0.1$, we have

$$\phi_{\epsilon, FM} \approx (2a'/\delta)\, J_1(\delta_d) \cos(\omega_{sc} t_d) \cos(\rho t_d/2) \qquad \text{(A49)}$$

Making use of the fact that $|J_1(\delta_d)|_{max} < 0.6$ as was done in Equation (A40), then the phase error for small a' is upper bounded by

$$\phi_{\epsilon,FM} < (0.075)\, a' \cos(\omega_{sc} t_d) \cos(\rho t_d/2) \tag{A50}$$

For a = 0.1,

$$|\phi_{\epsilon,FM}| < 0.0075 \text{ radians} = 0.43 \text{ degree}$$

Resultant Error in VOR Azimuth Measurement

The total error in bearing angle due to the single specular reflector is from Equations (A14) and A38) are given in exact form by

$$B_{\epsilon,VOR} = \left(\phi_{\epsilon,AM} - \phi_{\epsilon,FM}\right)_{VOR}$$

$$= -\left[\tan^{-1} \frac{a' \sin(\rho t_d + \theta_s - \theta_r)}{1 + a' \cos(\rho t_d + \theta_s - \theta_r)} - \tan^{-1} \frac{\gamma_1 \cos(\rho t_d/2)}{1 - \gamma_1 \sin(\rho t_d/2)}\right] \tag{A51}$$

For $a \leq 0.1$, $t_d \leq 0.1$ ms, and $|\theta_s - \theta_r| \geq 0.1$ radian, we have for the bearing error, the approximate expression

$$B_{\epsilon,VOR} \approx -\left[a' \sin(\theta_s - \theta_r) - 0.075\, a' \cos \omega_{sc} t_d\right]$$

$$\approx -a' \sin(\theta_s - \theta_r) \tag{A52}$$

where we have noted that the error in the FM reference tone is negligible compared with the bearing error in the information-bearing tone. For $\underline{a} = 0.1$, the maximum bearing error in the conventional VOR system $|B_{\epsilon, VOR}|_{max} \approx 5.7$ degrees, which constitutes a significant error in azimuth bearing.

Direct Plus Specular-Reflected Path:
Double Sideband Doppler VOR

The DSB doppler VOR (or DVOR) signal consists of a modulated RF carrier containing the following modulations:

a. A phase-reference tone at $\rho = 2\pi \times 30$ rad/sec amplitude modulated directly on the RF carrier.

b. A subcarrier at 9960 Hz, amplitude-modulated on the RF carrier, and sinusoidally frequency modulated (by induced doppler) by the azimuth-information signal $\delta \sin(\rho t + \theta_s)$ in which $\delta = 480/30 = 16$ rad, $\rho = 2\pi \times 30$ rad/sec, and θ_s is the azimuth angle to be measured.

c. Intelligence consisting of voice or a coded tone serving as a station identification signal, amplitude modulated onto the RF carrier.

Of main concern in this discussion is the effect of two-path propagation on the measurement of the azimuth angle, θ_s, in the VOR receiver. Accordingly, we express the DVOR signal as

$$e_{DVOR}(t) = E_s \left\{1 + m_1 \sin \rho t + m_2 \cos[\omega_{sc} t + \delta \sin(\rho t + \theta_s)]\right\} \cos(\omega_c t + \phi_c) \tag{A53}$$

where m_1 and m_2 are the amplitude modulation indices (= 30%), $\omega_{sc}/2\pi = 9960$ Hz, and $\cos(\omega_c t + \phi_c)$ represents the RF carrier.

The signal received by a distant aircraft in general consists of the direct signal plus reflections from a number of terrain features. In this analysis we consider only a single specular reflector to illustrate the method of computing the azimuth error introduced by multipath. This formulation, however, can be extended to the more general case of multiple reflectors, including scattering.

The single-reflector station model is again illustrated in Fig. A-2. The received DVOR signal consisting of the contributions of a direct path and of a second specular-reflection path, can be expressed in the form

$$\begin{aligned} e(t) = {} & E_s\{\cos[\omega_c t+\phi_c]+a\cos[\omega_c(t-t_d)+\phi_c]\} \\ & +m_1E_s\Big(\sin\rho t\cos[\omega_c t+\phi_c]+a\sin\rho(t-t_d)\cos[\omega_c(t-t_d)+\phi_c]\Big) \\ & +m_2E_s\Big(\cos[\omega_{sc}t+\delta\sin(\rho t+\theta_s)]\cos[\omega_c t+\phi_c] \\ & +a\cos\left[\omega_{sc}(t-t_d)+\delta\sin[\rho(t-t_d)+\theta_r]\right]\cos[\omega_c(t-t_d)+\phi_c]\Big) \end{aligned} \quad (A54)$$

where t_d represents the delay difference between the two paths, and a represents the relative amplitude ratio of reflected signal to direct signal. As above, we may view the envelope detector operation as product detection with a carrier whose phase is perfectly locked to the phase $\omega_c t+\phi_{c1}$ of the resultant carrier-phase reference for all sidebands. The result of envelope detection followed by lowpass filtering yields both the AM reference sidebands and the FM subcarrier sidebands.

AM Tone Detection

Envelope detection followed by lowpass filtering yields the AM reference tone

$$E_{AM}(t) = k_d m_1 E_s \cos\left[\phi_{c1}-\phi_c\right]\left[\sin\rho t + a'\sin\rho(t-t_d)\right] \quad (A55$$

where, as before

$$\cos[\phi_{c1}-\phi_c] = \frac{1+a\cos\omega_c t_d}{\sqrt{1+a^2+2a\cos\omega_c t_d}} \quad (A56)$$

$$a' \equiv a\,\frac{\cos[\omega_c t_d+\phi_{c1}-\phi_c]}{\cos(\phi_{c1}-\phi_c)} \quad (A57)$$

and k_d is a proportionality constant characteristic of the detector. The detected AM tone may be expressed in resultant form as

$$E_{AM}(t) = k_d m_1 E_s A_{AM}(a, t_d) \cdot \sin [\rho t + \phi_{\epsilon, AM}] \tag{A58}$$

where the real coefficient $A_{AM}(a, t_d)$ determines only the level of the detected AM tone. The phase error of the detected AM reference tone is given by

$$\phi_{\epsilon, AM} = -\tan^{-1} \frac{a' \sin \rho t_d}{1 + a' \cos \rho t_d} \quad \text{rad} \tag{A59}$$

Note that since $2\pi/\rho = 1/30$ sec, $\rho t_d \leq 0.1$ rad for all $t_d \lesssim 1$ msec, and we can write

$$\phi_{\epsilon, AM} \approx \frac{a'}{1+a'} \rho t_d \quad \text{rad for } t_d \lesssim 1 \text{ msec}. \tag{A60}$$

Since 1 msec corresponds to 300 kilometers of path-length difference, which is well beyond reasonable expectations, the above approximation is quite conservative. From the treatment of a' in the section entitled "Direct Plus Specular-Reflected Path: Conventional VOR," p. 80, we also note that the maximum value (and magnitude) of $\phi_{\epsilon, AM}$ occurs when $\cos \omega_c t = 1$, and this maximum is given by

$$\left|\phi_{\epsilon, AM}\right|_{max} \approx \frac{a}{1-a} \rho t_d, \quad \text{for } t_d = \frac{(2n+1)\pi}{\omega_c}, \; n = 0, 1, 2, \ldots$$

$$\approx (2n+1)\pi \left(\frac{a}{1-a}\right)(\rho/\omega_c), \; n = 0, 1, 2, \ldots \tag{A61}$$

For $t_d \leq 10$ μsec (corresponding to a maximum path-length difference of 3 kilometers);

$$\left|\phi_{\epsilon, AM}\right|_{max} \lesssim \frac{a}{1-a} \times 10^{-1} \text{ degrees} \tag{A62}$$

This peak value is 0.1 degree for a = 0.5. It must be remembered, however, that such a large value of a will normally be associated with values of t_d well below 10 μsec. In general, the larger values of $|a| < 1$ can be expected only with the smaller values of t_d (say 1 μsec or less), and the smaller a's with the larger t_d's.

FM Subcarrier Detection

The output of the receiver envelope detector also yields the frequency modulated subcarrier in which the bearing information, θ_s, appears as a phase shift of the modulating signal. The detected FM subcarrier is given by

$$E_{FM}(t) = k_d m_2 E_s \cos(\phi_{c1} - \phi_c) \left[\cos[\omega_{sc} t + \delta \sin(\rho t + \theta_s)] + a' \cos\left\{\omega_{sc}(t - t_d) + \delta \sin[\rho(t - t_d) + \theta_r]\right\}\right] \quad \text{(A63)}$$

where a' and $\cos(\phi_{c1} - \phi_c)$ are defined in Equations (A57) and (A56). The detected FM subcarrier can be expressed in resultant form as

$$E_{FM}(t) = k_d m_2 E_s A_{FM}(t) \cos[\omega_{sc} t + \delta \sin(\rho t + \theta_s) + \theta(t)] \quad \text{(A64)}$$

where

$$A_{FM}(t) = \cos(\phi_{c1} - \phi_c) \sqrt{1 + a'^2 + 2a' \cos \phi_F(t)} \quad \text{(A65)}$$

$$\theta(t) = -\tan^{-1} \frac{a' \sin \phi_F(t)}{1 + a' \cos \phi_F(t)} \quad \text{(A66)}$$

$$= \mathrm{Im}\left\{-\sum_{n=1}^{\infty} \frac{(-a')^n}{n} e^{jn\phi_F(t)}\right\}, \quad a'^2 < 1 \quad \text{(A67)}$$

and

$$\phi_F(t) = \omega_{sc} t_d + \delta \sin(\rho t + \theta_s) - \delta \sin[\rho(t-t_d) + \theta_r] \tag{A68}$$

We shall now derive an expression for the error $\phi_{\epsilon, FM}$ added to θ_s by the presence of the reflected signal. We proceed in a manner similar to the derivation of the phase error in the FM reference tone for the conventional VOR case. Thus, we first express Equation (A68) as

$$\phi(t) = \omega_{sc} t_d + \delta_d \sin(\rho t + \theta_s + \phi_d) \tag{A69}$$

where

$$\delta_d = 2\delta \sin\left(\frac{\rho t_d + \theta_s - \theta_r}{2}\right) \tag{A70}$$

$$\phi_d = -\frac{(\rho t_d + \theta_s - \theta_r - \pi)}{2} \tag{A71}$$

and, following the same procedure as before, derive the phase error which appears added to the bearing-information phase of the tone in the output of the discriminator. The bearing phase error on the FM tone in the DVOR system is given by

$$\phi_{\epsilon, FM} = -\tan^{-1} \frac{\gamma_1 \cos(\rho t_d + \theta_s - \theta_r)/2}{1 - \gamma_1 \sin(\rho t_d + \theta_s - \theta_r)/2} \tag{A72}$$

where γ_1 is given by Equation (A39) in which the function A_1, $n(a', \delta, \omega_{sc} t_d)$ is given by Equation (A37) with δ_d as defined in Equation (A70). An upper bounds on the FM phase error in Equation (A72) can be derived in a manner similar to the phase error bounds in Equation (A41). The result is

$$|\phi_{\epsilon, FM}| < \left| \tan^{-1} \frac{\cos[(\rho t_d + \theta_s - \theta_r)/2]}{\delta/[1.2 \ell n(1-a')] + \sin[(\rho t_d + \theta_s - \theta_r/2)]} \right| \tag{A73}$$

Special Case of Small a

The expression for the information-bearing tone phase error given by Equation (A72) can be simplified for the special case in which the amplitude ratio a is less than 0.1. The approximate phase error for $a \leq 0.1$ is given by

$$\phi_{\epsilon,FM} \approx (2a'/\delta)\, J_1(\delta_d) \cos(\omega_{sc} t_d) \cos(\rho t_d + \theta_s - \theta_r)/2 \quad \text{(A74)}$$

If we note that $|J_1(\delta_d)|_{max} < 0.6$, then for $\delta = 16$, we have

$$|J_1(\delta_d)|/\delta < 0.075 \quad \text{(A75)}$$

Thus, for $a = 0.1$,

$$|\phi_{\epsilon,FM}|_{max} \approx 0.0075 \text{ rad} = 0.43 \text{ degree} \quad \text{(A76)}$$

Resultant Error in DVOR Azimuth Measurement

The total error in bearing angle in the DVOR system due to the single specular reflector is, from Equation (A72) and Equation (A59), given in exact form by

$$B_{\epsilon,DVOR} = \left(\phi_{\epsilon,FM} - \phi_{\epsilon,AM}\right)_{DVOR}$$

$$= -\left[\tan^{-1} \frac{\gamma_1 \cos(\rho t_d + \theta_s - \theta_r)/2}{1 - \gamma_1 \sin(\rho t_d + \theta_s - \theta_r)/2} - \tan^{-1} \frac{a' \sin \rho t_d}{1 + a' \cos \rho t_d}\right] \quad \text{(A77)}$$

For $a \leq 0.1$, $t_d \leq 0.1$ msec and $|\theta_s - \theta_r| \geq 0.1$ radian, we can write

$$B_{\epsilon,\mathrm{DVOR}} \approx \left\{ \left[\frac{2}{\delta} J_1(\delta_d) \right] a' \cos \omega_{sc} t_d \cos \frac{(\theta_s - \theta_r)}{2} - \frac{a'}{1+a'} \rho t_d \right\} \tag{A78}$$

If we recall that $|J_1(\delta_d)| \leq 0.6$, $|a'| \leq a$ and $\delta = 16$, then

$$|B_{\epsilon,\mathrm{DVOR}}|_{\max} < (0.075)\, a$$

$$\approx 0.43 \text{ degree for } a = 0.1 \tag{A79}$$

These expressions and error bounds compare closely with those of Anderson and Flint[1]. In particular, Equation (A78), without the amplitude reference error term, is identical to Equation (A26) of the cited reference[2]. Moreover, the upper bound of 0.43 degree for $a = 0.1$ is exactly equal to the course scalloping error peaks in Fig. A-5 of Anderson and Flint[3].

Comparison of Bearing Errors of VOR and DVOR Systems

It is instructive to compare the bearing errors introduced by a specular reflection from one object for VOR and for DVOR systems operating in the same environment. The expressions derived above for $B_{\epsilon,\mathrm{VOR}}$ [Equations (A51) and (A52) and $B_{\epsilon,\mathrm{DVOR}}$ (Equations (A77) and (A78)] are repeated here for convenience:

---

1. S. R. Anderson and R. B. Flint, "The CAA Doppler Omnirange," Proc. IRE, vol. 47, pp. 808-821, May 1959.
2. Ibid., p. 813.
3. Ibid., p. 814.

$$B_{\epsilon,\,VOR} = -\left[\tan^{-1}\frac{a'\sin(\rho t_d+\theta_s-\theta_r)}{1+a'\cos(\rho t_d+\theta_s-\theta_r)} - \tan^{-1}\frac{\gamma_1\cos\rho t_d/2}{1-\gamma_1\sin\rho t_d/2}\right]$$

and

$$B_{\epsilon,\,DVOR} = -\left[\tan^{-1}\frac{\gamma_1\cos(\rho t_d+\theta_s-\theta_r)/2}{1-\gamma_1\sin(\rho t_d+\theta_s-\theta_r)/2} - \tan^{-1}\frac{a'\sin\rho t_d}{1+a'\cos\rho t_d}\right]$$

For $a \le 0.2$, $t_d < 0.1$ msec, and $|\theta_s - \theta_r| \ge 0.1$ rad, we have

$$B_{\epsilon,\,VOR} \approx a'\sin(\theta_s-\theta_r)$$

whence

$$|B_{\epsilon,\,VOR}|_{max} \approx a \text{ rad for } a \le 0.2$$

$$\approx 5.7 \text{ degrees for } a = 0.1$$

and

$$B_{\epsilon,\,DVOR} \approx -\gamma_1\cos(\theta_s-\theta_r)/2$$

$$\approx -\left\{\left(\frac{2}{\delta}\right)J_1\left[2\delta\sin(\theta_s-\theta_r)/2\right]\right\}a'\cos(\omega_{sc}t_d)\cos\frac{(\theta_s-\theta_r)}{2}$$

whence

$$|B_{\epsilon,\,DVOR}|_{max} \approx 0.43 \text{ degree, for } a = 0.1$$

Both error expressions are functions of $(\theta_s - \theta_r)$, the azimuth separation between aircraft and reflector headings. The VOR bearing error varies with $(\theta_s - \theta_r)$ and has peaks at $\theta_s - \theta_r - \pi/2$ and $3\pi/2$ radians and zeros at $(\theta_s - \theta_r) = 0$ or π radians. The DVOR bearing error also varies as a function of $(\theta_s - \theta_r)/2$. In addition, however, the bearing phase error in the DVOR experiences a much more rapid variation because of the influence of the Bessel coefficient. The argument of the Bessel coefficient $J_1(\delta_d)$ is a sinusoidal function of the bearing error. This argument is given by

$$\delta_d = 2\delta \sin \frac{(\theta_s - \theta_r)}{2}$$

where we have neglected ρt_d since $\rho t_d < 0.1$ degree for $t_d < 0.1$ msec. The value of the function $J_1(\delta_d)$ oscillates as a function of the azimuth bearing of the airplane relative to the fixed bearing of the reflector. The bearing error $B_{\epsilon, DVOR}$ also experiences these oscillations with the result that $B_{\epsilon, DVOR}$ goes to zero at the zeros of the function $J_1(\delta_d)$. The net effect of the fluctuation in the value of $J_1(\delta_d)$ produces a scalloped bearing error function. The periodicity of the scalloping, for constant airplane speed at a fixed location relative to the station, varies more rapidly than the bearing error of the conventional VOR system because the zeros of $J_1(\delta_d)$ occur more frequently, with changes in azimuth separation angle. Plots of the course scalloping caused by one reflecting object are compared in the paper by Anderson and Flint[4] where the effect of the rapid changes in $B_{\epsilon, DVOR}$ are clearly illustrated.

Direct comparison of the bearing expressions shows that the maximum bearing phase error due to reflection from a single object is significantly suppressed in the DVOR system relative to the maximum error in the conventional VOR system. In general, the bearing error in the DVOR due to reflections is suppressed to less than $1/\delta = 1/16$ of the bearing error in the VOR system, which confirms a theoretical result derived by Anderson and Flint[5]. There is no improvement, however, in the DVOR over the VOR system for azimuth headings of the aircraft coincident with the fixed heading of the reflector. In this special case the asymptotic bound on the bearing phase error with either system reduces to $\pm [\rho t_d + \theta_s - \theta_r]$ as $(\theta_s - \theta_r) \to 0$. Although there is no improvement by

---

[4] Ibid., p. 814, Figs. A-5 and A-6.

[5] Ibid., p. 814, Fig. A-7.

the DVOR system for this special case, the actual bearing phase error is nevertheless, below 0.5 degree regardless of which system is used.

The reduction in bearing error by the DVOR system is a direct advantage of conveying the information bearing tone in the form of an FM signal rather than an AM signal as in the conventional VOR system. With the information bearing tone frequency modulated onto the subcarrier, advantage is taken of the strong signal FM capture effect of the subcarrier discriminator to suppress the effects of the weaker reflected subcarrier. In the conventional VOR system, there is no suppression of the weaker component by the envelope detector and as a result the bearing phase error is directly proportional to $\underline{a}$ (for $a \leq 0.1$). The error reduction factor of the DVOR system relative to the VOR system is embodied in the coefficient

$$\left(\frac{2}{\delta}\right) J_1 \left[2\delta \sin (\theta_s - \theta_r)/2\right] < 0.075$$

It is important to note that the DVOR system will also show significant improvement over the VOR in cases where the reflected signal is very strong approaching the direct signal in strength. In such a case the error would be intolerable in a conventional system. Assuming no distortion in the envelope detector which delivers the FM subcarrier, then the bearing phase error in the DVOR will always be proportional to $\underline{a}$ multiplied by a suppression factor equivalent to the above coefficient. Suppression of the interfering reflections will diminish however as $\underline{a}$ approaches unity depending upon the capture performance of the FM detector. For $\underline{a} \lesssim 1$, the distortion and resulting bearing phase error will be large and for $a>1$ the FM detector will suppress the true information bearing tone and capture the reflected tone. Such situations are unlikely in practice since the reflected signal usually experiences greater path attenuation because of the larger distances traveled relative to the direct path. Moreover, the properties of the reflecting surface will in general have a reflection coefficient less than one and will unavoidably attenuate the reflected signal so that the received reflected component will usually be weaker than the direct-path signal.

Direct Plus Specular-Reflected Path: Single Sideband Doppler VOR

The original conception of the Doppler VOR system was such that only one sideband was radiated by the station. This single sideband is displaced from the VHF carrier by 9960 Hz and is sinusoidally frequency modulated (by induced doppler) by the azimuth information signal

$\delta \sin(\rho t + \theta_s)$ in which $\delta = 480/16 = 30$ rad, $\rho = 2\pi \times 30$ rad/sec and θ_s is the azimuth angle to be measured. The ratio of the single sideband level to the carrier is $m_2 = 0.3$. The AM reference tone and station intelligence is identical to that described for the DSB-DVOR system. Of main concern in this discussion is the effect of two-path propagation on the measurement of the azimuth angle, θ_s, in the VOR receiver. We express the single sideband DVOR (SVDOR) signal as

$$e_{SDVOR}(t) = E_s \{1 + m_1 \sin \rho t\} \cos(\omega_c t + \phi_c) + m_2 E_s \cos\{(\omega_c + \omega_{sc})t + \delta \sin(\rho t + \theta_s) + \phi_c\}$$

where m_1 and m_2 are the amplitude modulation indices (= 30%), $\omega_{sc} = 2\pi \times 9960$ Hz and $\cos(\omega_c t + \phi_c)$ represents the RF carrier. We have arbitrarily chosen an upper sideband representation; the analysis, however, does not depend upon which sideband is selected.

The received SDVOR signal consisting of the contributions of a direct path and of a second specular-reflection path, can be expressed in the form

$$\begin{aligned} e(t) = {} & E_s\{\cos[\omega_c t + \phi_c] + a \cos[\omega_c(t - t_d) + \phi_c]\} \\ & + m_1 E_s\{\sin \rho t \cos(\omega_c t + \phi_c) + a \sin \rho(t - t_d) \cos[\omega_c(t - t_d) + \phi_c]\} \\ & + m_2 E_s\{\cos[(\omega_c + \omega_{sc})t + \delta \sin(\rho t + \theta_s) + \phi_c] \\ & \quad + a \cos[(\omega_c + \omega_{sc})(t - t_d) + \delta \sin[\rho(t - t_d) + \theta_r] + \phi_c]\} \end{aligned}$$

where t_d represents the delay difference between the two paths, and $\underline{a}$ represents the relative amplitude ratio of reflected signal to direct signal. As in section entitled "Direct Plus Specular-Reflected Path: Conventional VOR," p. 80, we may view the envelope detector of the AM reference sidebands as product detection with a carrier whose phase is perfectly locked to the phase $\omega_c t + \phi_{c1}$ of the resultant carrier-phase reference for the AM sidebands.

AM Tone Detection

The result of envelope detection followed by lowpass filtering yields the AM reference sidebands, and the error analysis is identical to that of the DSB-DVOR case. The effect of the reflected signal produces the same error in the phase of the reference tone. This error was derived in section entitled "Direct Plus Specular-Reflected Path: Double Sideband Doppler VOR," Equation (A59), p. 97, and is repeated here.

$$\phi_{\epsilon, AM} = - \tan^{-1} \frac{a' \sin \rho t_d}{1 + a' \cos \rho t_d} \quad \text{rad}$$

Error in the phase of the information bearing tone is therefore given by Equation (A72) and repeated here.

$$\phi_{\epsilon, FM} = - \tan^{-1} \frac{\gamma_1 \cos (\rho t_d + \theta_s - \theta_r)/2}{1 - \gamma_1 \sin (\rho t_d + \theta_s - \theta_r)/2}$$

The bearing error for the SSB-DVOR system is thus given by Equation (A77) and repeated here.

$$B_{\epsilon, DVOR} = \left(\phi_{\epsilon, FM} - \phi_{\epsilon, AM} \right)_{DVOR}$$

$$= - \left[\tan^{-1} \frac{\gamma_1 \cos (\rho t_d + \theta_s - \theta_r)/2}{1 - \gamma_1 \sin (\rho t_d + \theta_s - \theta_r)/2} \right.$$

$$\left. - \tan^{-1} \frac{a' \sin \rho t_d}{1 + a' \cos \rho t_d} \right]$$

and for $a \leq 0.1$

$$\left| B_{\epsilon, SDVOR} \right|_{max} \approx 0.43 \text{ degree.}$$

APPENDIX B

PSEUDORANDOM RANGING ERROR

In this appendix the tracking error of the scheme shown in Fig. 2 is computed. The received signal is of the form

$$\sqrt{2S}\quad s(t)\cos\omega_c t + n(t) \tag{B1}$$

The signal power is S (the value of S(t) is either + 1 or -1 in each baud) and the double-sided spectral density of n(t) is $N_o/2$. If this is converted coherently to baseband, the resulting baseband signal is

$$\sqrt{S}\ \ s(t)\ \ \cos\omega_c t + n'(t) \tag{B2}$$

The density of the new noise process n'(t) is also $N_o/2$. The signal of Equation (B2) forms the input to the baseband receiver structure of Fig. 2.

The two identical integrators of Fig. 2 are linear filter with the impulse response:

$$h_1(\tau) = \begin{cases} \dfrac{1}{T_{int}} & O < \tau < T \\ O & \text{elsewhere} \end{cases} \tag{B3}$$

The noise voltage at the output of the summer is given by

$$n_{out} = \left[n'(t)\ s(t+\Theta)\right] * h_1(t)$$

$$- \left[n'(t)\ s\,(t+\Theta-2T)\right] * h_1(t) \tag{B4}$$

where the symbol * implies convolution.

Since the operation of convolution is linear this may be rewritten as

$$n_{out} = \left\{ n'(t) \left[S(t + \Delta) - S(t + \Delta - 2T) \right] \right\} * h_1(t) \quad \text{(B5)}$$

Notice that the quantity in the square bracket assumes the values 2, 0, and -2 with relative frequencies 1/4, 1/2, and 1/4 (respectively) due to the pseudorandom nature of the waveform s (t). Only the rms value of n_{out} is of concern, so the factor in question will be replaced by its rms value which is

$$\sqrt{(2)^2\left(\frac{1}{4}\right) + (0)^2\left(\frac{1}{2}\right) + (-2)^2\left(\frac{1}{4}\right)} = \sqrt{2} \quad \text{(B6)}$$

This manipulation is justified by the fact that $[s(t + \theta) - s(t + \theta - 2T)]$ is statistically independent of n'(t). Thus, the output noise voltage can be written as

$$n_{out} = \sqrt{2}\ n'(t) * h_1(t) \quad \text{(B7)}$$

and the mean square noise output is given by

$$\overline{n_{out}^2} = 2\left(\frac{N_o}{2}\right)\int_{-\infty}^{\infty} |H_1(\omega)|^2 \frac{d\omega}{2\pi}$$

$$= N_o \int_{-\infty}^{\infty} |h_1(t)|^2 dt$$

$$= \frac{N_o}{T_{int}} \quad \text{(B8)}$$

Referring to Fig. 5, timing error may be related to noise voltage by the slope of the control characteristic at the point labeled $\theta = T$. By inspection this slope is

$$\frac{T}{\sqrt{S}} \tag{B9}$$

Combined with Equation (B8) this gives the rms timing error

$$\sqrt{\overline{(\Delta t)^2}} = \frac{T}{\sqrt{T_{int} \frac{S}{N_o}}} \tag{B10}$$

Converted to range error this gives the result of Equation (B4).

APPENDIX C

BIT SYNCHRONIZATION TIMING ERROR

In this Appendix we analyze the performance of the symbol synchronization loop of Fig. 11 following Viterbi, we write the output of the squarer as

$$\left[s(t) + n(t) \right]^2 = s^2(t) + \nu(t) \tag{C1}$$

where

$$\nu(t) = 2s(t)\, n(t) + n^2(t) \tag{C2}$$

and n(t) is the noise voltage at the squarer input. Viterbi shows that the spectrum of $\nu(t)$ is given by

$$S_\nu(\omega) = 4 \int_{-\infty}^{\infty} S_n(\omega - \zeta)\, S_s(\zeta)\, \frac{d\zeta}{2\pi} + 2 \int_{-\infty}^{\infty} S_n(\omega - \zeta)\, S_n(\zeta)\, \frac{d\zeta}{2\pi} \tag{C3}$$

The noise spectrum $S_n(\omega)$ is flat over a bandwidth ω equal to the bandwidth of the bandpass filter preceding the squarer. Therefore, from (C3), the spectrum $S_\nu(\omega)$ will be uniform over the much narrower bandwidth B_ν of the phase-locked loop. Therefore, the effective one-sided noise density at the loop input is equal to twice the noise density at the loop center frequency:

$$N_o' = 2S_\nu\left(\frac{2\pi}{T}\right)$$

$$= 8 \int_{-\infty}^{\infty} S_n\left(\frac{2\pi}{T} - \zeta\right) S_s(\zeta)\, \frac{d\zeta}{2\pi} +$$

$$4 \int_{-\infty}^{\infty} S_n \left(\frac{2\pi}{T} - \zeta\right) S_n (\zeta) \frac{d\zeta}{2\pi} \tag{C4}$$

The spectrum $S_n(\omega)$ is centered about $\pm\omega_c$, and has height $N_o/2$; carrying out these integrals gives

$$N_o' = 4 N_o A^2 + 2 N_o^2 \left(W - \frac{1}{T}\right) \tag{C5}$$

where A^2 is the signal amplitude and W is the bandwidth of the bandpass filter in Fig. 11. The first term in this result is twice the corresponding result in Viterbi because both the positive and negative halves of the signal spectrum fall into the region of integration. The second term is also doubled but is reduced by 1/T because of the offset $2\pi/T$. From Viterbi, the mean-square phase error of the tracking loop is given by

$$\sigma_{2\phi}^2 = \frac{N_o' B_L}{A^4/2} \tag{C6}$$

where B_L is the loop bandwidth and $A^4/2$ is the power of the symbol frequency component after the squarer. Combining these two results and letting W = 6/T for negligible signal distortion, gives

$$\sigma_{2\phi}^2 = \frac{8N_o B_L}{A^2} + \frac{20N_o^2 B_L}{A^4 T} \tag{C7}$$

or

$$\sigma_{\phi}^2 = \frac{2N_o B_L}{A^2} \left(1 + \frac{5N_o}{2A^2 T}\right) \tag{C8}$$

This may be converted to units of time with the relationship $\phi = \omega_{BR} t$ giving

$$\sigma_t^2 = \left(\frac{T}{2\pi}\right)^2 \frac{2N_o B_L}{A^2} \left(1 + \frac{5N_o}{2A^2 T}\right) \tag{C9}$$

A suitable bit error probability is insured by the requirement

$$\frac{E}{N_o} = \frac{A^2 T}{N_o} = 10 \tag{C10}$$

Then (C9) becomes

$$\sigma_t^2 = \frac{T^3 B_L}{(4\pi)^2} \tag{C11}$$

or,

$$\sigma_t = \frac{T}{4\pi}\sqrt{T B_L} \tag{C12}$$

APPENDIX D

NEW TECHNOLOGY

After a diligent review of the work performed under this contract, no new innovation, discovery, improvement or invention was made.

DISTRIBUTION LIST

1. NASA/Electronics Research Center
575 Technology Square
Cambridge, Massachusetts 02139

AT/Technical Information Branch (20 + 1 reproducible)
Technical Monitor (15)
T/Technology Utilization (1)
M/Patent Counsel (1)

2. NASA Scientific and Technical Information Facility (3)
P.O. Box 33
College Park, Maryland 20740

3. NASA Lewis Research Center (2)
21000 Brookpark Road
Cleveland, Ohio 44135

4. National Aeronautics and Space Administration (1)
Washington, D.C. 20546
Attention: William H. Woodward (RN)

5. T.K. Speer, Department Engineer (1)
Advanced Electronics Systems Department
Lockheed-Georgia Company
Marietta, Georgia 30060